数学学科专题史丛书

代数数论简史

Concise history of algebraic number theory

冯克勤 著

湖南教育出版社

代数数论简史

Concise history of algebraic number theory

内容简介

代数数论是数论的一个重要分支。经典代数数论研究代数数域的数论和代数性质，起源于德国数学家高斯和库默尔，距今有两百年历史。近代和现代的代数数论则与几何、分析、代数相互交织，构成当前纯粹数学活跃的研究领域。1994年怀尔斯证明费马猜想是这领域发展的一个重要标志。另一方面，1960年以来由于数字计算机和网络通信技术的飞速进步，代数数论在计算机科学和信息科学中得到重要应用。

本书较为通俗地介绍了代数数论的发展轮廓，讲述各历史时期数学家们创造的主要思想和方法以及代数数论的重要成就。还扼要介绍了代数数论的某些重要应用。

本书的读者范围为大学生、研究生、大学数学教师和数学研究人员，也包括从事信息领域工作的工程师和其他数学爱好者。

ISBN 7-5355-3741-3
9 787535 537416 >

代数数论简史

Concise history of algebraic number theory

冯克勤 著

湖南教育出版社

图书在版编目(CIP)数据

代数数论简史 / 冯克勤著. —长沙：湖南教育出版社，2001

(数学学科专题史丛书)

Ⅰ.代… Ⅱ.冯… Ⅲ.代数数论—数学史 Ⅳ.0156.2-09

中国版本图书馆 CIP 数据核字(2001)第 07287 号

代数数论简史

冯克勤 著

责任编辑：孟实华

湖南教育出版社出版发行(长沙市韶山北路 643 号)

湖南省新华书店经销 湖南广播电视大学印刷厂印刷

850×1168 32 开 印张：8.875 字数：228000

2002 年 7 月第 1 版 2002 年 7 月第 1 次印刷

ISBN 7-5355-3741-3/G·3736

定价：17.00 元

前言

希尔伯特在著名的《数论报告》(1897年)前言的开头便指出："数论是最古老的数学分支之一.人类的心灵很早就注意到自然数某些相当深刻的性质.但是数论作为一门独立和系统的科学则完全是现代的成就."

人类在有文字之前便在生产和生活实践中建立了某些整数的概念.三千多年前,东方文明古国(埃及、巴比伦、印度、中国)开始研究整数各种性质和方程(组)的整数解(和有理数解),由此产生了数论.古代中国对数论有许多杰出贡献(勾股定理和勾股数、中国剩余定理等).古代东方的数论具有鲜明的实践、直觉和算法特点,而古希腊的数论(公元前6世纪到公元3世纪)则具有理性和思辨的特征("万物皆数"、惟一因子分解、素数的无限性、方程 $x^2+y^2=z^2$ 的全部整数解……).数论的近代发展是从17世纪开始的,数论中心在法国(除了欧拉之外,17和18世纪主要数论学家费马、勒让德、拉格朗日、拉普拉斯等均为法国人).希尔伯特所说的"现代"便是从这时开始的.19世纪数论取得了重大进步,集中体现为研究方法的创新、深刻的代数工具和解析工具引入到数论问题的研究中,产生了两个重要的数论分支——代数数论和解析数论.如果把高斯在《数论探究》(1801)一书中对二元二次型和二次域的研究作为代数数论的开端,那么代数数论至今恰好有二百年的历史.

从高斯的《数论探究》到希尔伯特的《数论报告》,整个19世纪是经典代数数论产生、建立、发展和定型的时期.19世纪初期高斯通过二元二次型的分类和表示整数问题对二次域做了深入研究.19世纪中期库默尔在研究费马猜想的过程中,对分圆域做了系统

的理论工作，在这个基础上，狄德金把他们的思想加以确切化和代数化，建立了一般代数数域上的基本概念和结果. 到了 19 世纪末，由希尔伯特作了更全面和系统的总结. 经典代数数论是研究代数数域和它的代数整数环的代数和算术性质的一个数论分支. 代数数域即是有理数域的任意有限次扩域，经典代数数论的主要课题是研究整基、素数在代数数域中的素理想分解、理想类群和类数以及代数整数环的单位群结构(基本单位系).

代数数论在 20 世纪得到全面的发展. 1900 年希尔伯特在第二届世界数学家大会(巴黎)上提出了 23 个数学问题. 对于其中 4 个代数数论问题的研究在很大程度上促进了 20 世纪代数数论的进步. 这个时期代数数论研究中出现了许多新思想和方法(赋值和局部域理论、局部—整体原则、p-adic 分析、几何和解析方法的引进等). 与代数几何、复分析、近世代数等结合在一起，产生了一系列重要的研究领域[类域论、模形式理论、代数曲线(特别是椭圆曲线)的算术理论、分圆域近代理论等]. 20 世纪(特别是后 30 年)也是代数数论大丰收的时代，它集中表现为三个费尔兹奖获得者和两个沃尔夫奖获得者的工作. 1973 年德林(P. Deligne)证明了高维韦依猜想和 1983 年法廷斯(Faltings)证明了莫代尔猜想，推动算术几何的发展，这两项工作分别获得 1978 年和 1986 年的费尔兹奖. 而 1990 年获得费尔兹奖的德林费尔德(Drinfeld)证明了函数域上 2 维局部朗兰兹猜想，是宏大的朗兰兹纲领的最重要突破. 朗兰兹(Langlands)是 1996 年沃尔夫奖得主，他于 20 世纪后期提出的一系列猜想是用群的无限维表示理论研究数论，把代数数论和群上调和分析，微分几何和微分方程，李群和李代数等诸多数学分支交织在一起，被人们与克莱因(Klein)的爱尔朗根纲领(用群论研究几何)相提并论. 与朗兰兹同时获得沃尔夫奖的怀尔斯(A. Wiles)于 1994 年证明了具有 350 多年历史的费马猜想，则藉助于现代代数数论诸多方面的最新成就. 代数数论在 20 世纪后期不仅取得丰富的理论成果，而且在计算机科学和信息工程领域得到重

要应用，它正在同时扮演着皇后和仆人的角色.

这本书的目的是试图简要地介绍代数数论二百年的发展途径.这不是处处引经据典的考古性历史著作，也不是力求全面的百科全书，它只是沿着历史的线索讲述代数数论的主要思想、方法、成就和一些重大的事件.书中的取材和观点不可避免地伴有作者的偏好和片面性，作者希望能厚今薄古，但是近代和现代的代数数论内容如此丰富，所涉及的相关数学领域如此之广，研究成果如此博大精深，使作者深感心有余而力不足.对于想学习代数数论和有志于从事代数数论研究的读者(特别是年轻人)，以及打算以代数数论为工具的技术领域的专家们，希望本书对他们有所帮助，对代数数论有一个整体性的宏观了解.

本书共分五章.第一章讲述代数数论的前身，试图阐明古代数论特别是17和18世纪的数论发展如何导致代数数论的诞生.第二章讲述代数数论的前一百年(从高斯到希尔伯特)，即19世纪经典代数数论的产生、建立、发展和定型.第三章讲述20世纪前70年近代代数数论丰富的内容，第四章讲述20世代后30年现代代数数论的重大理论成就.第五章讲述代数数论于20世纪后半期在信息科学领域的某些重要应用.

由于学识和资料所限，书中内容和观点会有不妥之处，欢迎批评指正.作者感谢国家自然科学基金委员会和湖南教育出版社对出版此书给予的支持.本书的大部分是在1999年夏天访问意大利底里亚斯特国际理论物理中心(ICTP，Trieste)期间写成的，作者感谢ICTP所提供的图书资料和舒适安静的环境.

冯克勤

1999年11月

目　　录

第一章　代数数论的前身

§1.1　古代数论

人类在有文字历史之前就有了一些自然数的概念.在生产和社会生活实践中,从两只手、两头羊之中抽象出数"2"的概念,是人类理性的巨大飞跃.在文字发明之后,不同地区的人们相继出现了各种不同形式的数字(表达数的文字).随着社会生产的发展,发明了表达大整数的计数方法(各种进位制和位值制)以及数的运算,探讨整数的各种性质和方程(组)的整数解(和有理数解),这就产生了数论.数论是最古老的数学学科之一,它至少有三千年的历史.

数论产生于东方文明古国(埃及、巴比伦、印度、中国).古代数论研究最初是与几何学交织在一起的.中国最早的数学著作《周髀算经》(约公元 2 世纪)中记载,西周人商高知道直角三角形三边长满足方程 $x^2+y^2=z^2$,并且给出此方程一组正整数解 $(x,y,z)=(3,4,5)$(即勾三股四弦五).印度人在公元前也知道此方程的特殊解.据有人考证,巴比伦人在公元前 18 世纪就知道此方程的 15 个正整数解,很可能他们已经知道一般解 $(x,y,z)=(2uv,u^2-v^2,u^2+v^2)$,例如他们知道解(13500,12709,18541)(即取 $u=125,v=54$).西方数学把 $x^2+y^2=z^2$ 的正整数解叫作毕达哥拉斯数,事实

上毕达哥拉斯是向巴比伦人学习了此方程的求解方法. 古代人类还对其他形式的代数方程寻求整数解或有理数解(详见 Dickson:《History of Number Theory》,1971 年版). 其中包括后来被欧拉称为 Pell 方程的 $x^2-dy^2=1$. 印度数学家 Bhaskara(1114—1185)得到方程 $x^2-61y^2=1$ 的最小正整数解 $(x,y)=(1766319049, 226153980)$,他用的方法相当于现在的连分数方法. 古代中国对整数同余性质有深刻的认识. 在《孙子算经》中载有"物不知数"问题,给出一次同余方程组的解法. 这种方法在近代数学中已被推广成非常一般的代数形式,但仍被世人称为"中国剩余定理".

古代东方的数论具有鲜明的实践、直觉和算法特点. 而古希腊的数论(约公元前 6 世纪至公元 3 世纪)则具有理性和思辨的特征. 毕达哥拉斯(Pythagoras,约公元前 570 年～公元前 500 年)学派的格言是"万物皆数",把数量关系看作是人认识各种事物的最本质手段. 他们认为事物是"可公度"的,即可通过整数或整数之间比例(有理数)测量来认识事物. 但是毕达哥拉斯的学生通过理性思考发现正方形一边和对角线不是可公度的(即$\sqrt{2}$不是有理数). 他们所崇拜的正五边形的边和对角线也不可公度(即$\frac{\sqrt{5}-1}{2}$也不是有理数). 他们对整数的性质进行着迷地研究. 比如发现了前几个偶完全数 6,28 和 496(正整数 n 叫作完全数,是指 n 的所有正的真因子之和等于 n). 欧几里德证明了:若 $p=2^m-1$是素数,则 $n=2^{m-1}(2^m-1)$必为完全数. 一直到两千年之后欧拉才证明了:所有偶完全数都具有这种形式. 而奇完全数是否存在,至今仍是个谜.

古希腊数论成就集中反映在欧几里德(Euclid,公元前 330 年～前 275 年)于公元前 3 世纪所写的《几何原本》(Elements)一书中. 这本书共 13 卷,其中有 5 卷讲数论. 从以下主要内容可以看出古希腊数论的深刻性.

(1) 欧氏除法算式(即带余除法):设 a 和 b 为正整数,$a>b$.

则存在惟一的整数 q 和 r 使得

$$a=qb+r,$$

其中 $q\geqslant 1$，而 $0\leqslant r<b$. 利用这个结果，欧几里德还给出了求最大公因子的辗转相除算法.

(2) 若素数 p 除尽整数 a 和 b 的乘积 ab，则 p 除尽 a 或 b. 利用这个性质，欧几里德证明了数论的最基本和重要的性质：

(3) (算术基本定理)每个大于 1 的整数 n 不计因子次序都可惟一地表示成有限个素数之乘积.

事实上，欧几里德对于惟一性只陈述和证明了较弱结果：如果 n 没有平方因子，把 n 的素因子从小到大排列 $p_1<p_2<\cdots<p_r$，则分解式 $n=p_1p_2\cdots p_r$ 是惟一的. 算术基本定理的完整证明是高斯于 1801 年给出的. 但是从欧几里德时代开始，人们无疑地使用着算术基本定理这块数论的基石. 我们将会看到：整数的推广并考查在扩大的整数环中惟一因子分解性质是否成立，是产生代数数论的一个重要因素.

(4) 利用算术基本定理. 欧几里德事实上给出了方程 $x^2+y^2=z^2$ 的全部正整数解：

$$x=a(m^2-n^2),\quad y=2amn, z=a(m^2+n^2).$$

其中 a,m,n 均为正整数($m>n$). 但是按照 Dickson 的考证，完整的证明是 1738 年由 C. A. Koerbero 给出的. 我们这里只讲证明中最关键的步骤.

(A) 首先可设 x,y,z 两两互素，这种正整数解叫**本原解**，而其他正整数解都可由某个本原解乘以同一个正整数而得到.

(B) 对于本原解 (x,y,z)，易知 z 为奇数，而 x 和 y 一奇一偶. 不妨设 y 奇 x 偶. 于是方程可改写为

$$x^2=z^2-y^2=(z-y)(z+y).$$

这一步虽然简单，但是很关键，因为它把多项式 z^2-y^2 写成乘积形式 $(z-y)(z+y)$. 将这个等式改写为

$$\left(\frac{x}{2}\right)^2=\frac{z-y}{2}\cdot\frac{z+y}{2},$$

其中$\frac{x}{2}$,$\frac{x\pm y}{2}$均为正整数.

(C) 再证$\frac{z-y}{2}$和$\frac{z+y}{2}$互素.由算术基本定理可知两个互素正整数之积为平方数,则每个正整数都是平方数.于是$\frac{z+y}{2}=m^2$,$\frac{z-y}{2}=n^2$,从而$z=m^2+n^2$,$y=m^2-n^2$,而$x=2mn$(其中m和n是正整数,$m>n$).这本质上给出方程$x^2+y^2=z^2$的所有正整数解.

在研究其他不定方程整数解时,上述证明采用的方法遇到了困难.对于步骤(B),有些方程往往不能化成乘积表达形式,这需要扩大整数范围.对于步骤(C),在扩大的整数环中不再有惟一因子分解性质,这是更严重的困难.对这一困难的深入思考导致代数数论的进步.

(5) 欧几里德证明了素数有无限多个.

证明用反证法:

如果素数只有有限多个$p_1,p_2,\cdots,p_r$.根据算术基本定理,$n=p_1p_2\cdots p_r+1$为有限个素数的乘积.但是素数只有$p_1,\cdots,p_r$,而它们均不是n的因子,这就导致矛盾.

这可能是世界上第一个采用反证法证明无限性的命题.

古希腊另一本重要数论著作是丢番都(Diophantus,约公元250年)的《算术》(Arithmetica).书中收集300多个数论问题,给出各种形式的二元和三元代数方程(组)(次数$\leqslant 3$)寻求整数解和有理数解的方法.其中包括方程$x^2+y^2=z^2$,Pell方程$x^2-dy^2=1$和方程$x^2+k=y^3$(对某些特殊值d和k).这是世界上第一本离开几何而专门讲述数论的著作.我们即将看到,这本著作对于代数数论的发展具有特殊的意义.



图 1　埃及纪念碑上人们用手指计算的图画(公元前 26 世纪)

§1.2　17 和 18 世纪的数论

世界文明中心逐渐转移到欧洲. 在经过中世纪黑暗时代之后, 欧洲于 15 和 16 世纪为文艺复兴时代. 基于航海、建筑、绘画和雕塑等发展和需要, 画法几何学在这一时期取得很大进步, 而数论没有重要的成就. 17 世纪为欧洲产业革命时期. 反映连续性的数学——微积分和将几何加以代数化的解析几何产生和发展, 数学分析在后来三百年内占据数学的主导地位. 而数论在 17 世纪出现了一位杰出的传奇性人物: 费马. 18 世纪数论取得很大发展, 当时的世界数论中心在法国. 除了费马之外, 法国数论学家还有拉格朗日(Lagrange, 1736—1813), 勒让得(Legendre, 1752—1833)等, 这个时期大数论学家中只有欧拉(Euler, 1707—1783)不是法国人.

费马(Pierre de Fermat, 1601—1665)一生的主要职业是法国南部小城 Toulouse 的律师和公务员. 数学是他的业余爱好. 他与 Pascal 合作研究过概率论, 并且是解析几何的创始人之一(另一个为哲学家笛卡儿 Descartes). 但是他对数学的最大贡献是数论. 他

一生从未写过一本著作，只写过一篇关于解析几何的文章．他的数论思想都写在读书评注和与朋友的通信中，这些工作在他去世后由他的儿子整理出版．



图 2　费马(Fermat，1601—1665)

1637 年，费马在阅读丢番都《算术》一书的拉丁文版本时，在论述方程 $x^2+y^2=z^2$ 的那一页写了一个评注：

"任何立方数都不是两个立方数之和.任何(整)数的四次方都不是两个四次方数之和.一般地,任何一个更高次方均不是两个同次方幂之和.我有一个确实奇妙的证明,但是地方太小写不下."

也就是说,费马提出如下的猜想:对每个正整数 $n\geqslant3$,方程 $x^n+y^n=z^n$ 没有正整数解.这也相当于说,此方程没有非零整数解(即 x,y,z 均为非零整数).

事实上,人们只看到费马对于 $n=4$ 给出的证明.他证明了更好的结果:方程 $x^4+y^4=z^2$,没有正整数解.他的证明用反证法.

设方程 $x^4+y^4=z^2$ 存在正整数解,则存在正整数解 (x,y,z) 使得 z 最小.这时,x,y,z 必然两两互素.因为若 x,y,z 当中任何两个有正整数公因子 $a>1$,则另一个也被 a 除尽.令 $x=ax'$,$y=ay'$,则 x' 和 y' 为正整数,而方程变为 $a^4(x'^4+y'^4)=z^2$,这表明 z 可被 a^2 除尽.令 $z=a^2z'$,则 $(x'^4+y'^4)=y'^2$ 即 (x',y',z') 又是原方程一组正整数解,而正整数 $z'=z/a^2$ 小于 z,与假设 z 最小矛盾.于是 x,y,z 两两互素.将方程写成 $(x^2)^2+(y^2)^2=z^2$,可知 (x^2,y^2,z) 是 $X^2+Y^2=Z^2$ 的本原解.由欧几里德结果知(不妨设 x 为偶数)

$$x^2=2uv,\ y^2=u^2-v^2,\ z=u^2+v^2,$$

其中 u,v 为互素的正整数 $(u>v)$.由第二个方程知 $u^2=y^2+v^2$,从而 u 为奇数而 v 为偶数(否则 x,y,z 均为偶数与假设矛盾).于是又有

$$v=2st,\ y=s^2-t^2,\ u=s^2+t^2,$$

其中 s,t 为整数.这时 $x^2=2uv=4st(s^2+t^2)$.由于 s 和 t 互素,从而它们与 s^2+t^2 也互素.上式表明 s,t 和 (s^2+t^2) 均是平方数,即

$$s=x_1^2,\ t=y_1^2,\ s^2+t^2=z_1^2.$$

于是 $x_1^4+y_1^4=z_1^2$,但是 z_1 是比 z 小的正整数.这就与 z 的最小性矛盾.

这一矛盾表明 $x^4+y^4=z^2$ 不存在正整数解(从而 $x^4+y^4=z^4$ 也是如此).

费马对他的证明颇为得意，起了名字叫“无穷下降”法. 即若 (x,y,z) 是某方程的一组正整数解，如果由此可推出该方程另一组正整数解 (x',y',z')，使得 $z'<z$. 继续下去便推出矛盾，因为正整数 z 不能无限地递降. 这就表明该方程没有正整数解. 或许费马认为用这种方法可以证明任意 $n\geqslant 3$ 的情形，但事实远不是那样简单. 过了一百多年，欧拉对费马这个猜想才（基本上）证明了 $n=3$ 的情形. 但我们今后将会看到，无穷下降法是数论研究的有力工具.

数论不是 17 世纪的热门数学，费马研究数论没有竞争对手，这也许是他没有发表工作的原因之一. 他提出许多数论猜想，这些猜想是根据一些计算实例和他对数论的直觉，均没有给出证明. 下面是他的一部分猜想.

（1） 1640 年，费马在给 Frenicle 的信中说：若整数 a 不被素数 p 除尽，则 a^{p-1} 被 p 除余 1. 用高斯发明的同余式语言，费马的论断可写成 $a^{p-1}\equiv 1 \pmod{p}$. 这个论断是在 18 世纪由拉格朗日证明的，后人称为费马小定理.

而欧拉又把它推广成更一般的结果：若整数 a 与正整数 $m\geqslant 2$ 互素，则 $a^{\varphi(m)}\equiv 1 \pmod{m}$，其 $\varphi(m)$ 表示 $1,2,\cdots,m$ 当中与 m 互素的正整数个数. 后人把 $\varphi(m)$ 叫作欧拉函数. 当 m 为素数 p 时，$\varphi(p)=p-1$. 从而费马小定理是上述欧拉定理的特殊情形.

（2） 对于正整数 m，如果 m 有奇素数因子 p，易知 2^m+1 可被 $2^{\frac{m}{p}}+1$ 除尽. 所以若 2^m+1 为素数，则 m 只能是 2 的方幂. 于是费马考虑数 $F_n=2^{2^n}+1$. 他算出 $F_0=3, F_1=5, F_2=17, F_3=257$ 和 $F_4=65\ 537$ 都是素数. 于是他猜想：对所有 $n\geqslant 0$，F_n 均是素数（后人称之为费马素数）. 但是欧拉巧妙地发现了 641 为 F_5 的因子，从而推翻了费马这个猜想. 事实上，一直到今天除了费马给出的前 5 个之外，人们还没有发现新的费马素数，甚至倾向于：当 $n\geqslant 5$ 时，F_n 均不是素数. 目前已经算出了 $F_n(n\leqslant 10)$ 的素因子分解式从 F_7

到 F_{10} 的素因子分解式都是利用最先进计算机和基于高深代数数论知识研制出的大数分解算法而得到的. 我们在本书最后一章将会看到,近年来人们热衷于分解 F_n 的动机来自保密通信.

(3) 1657 年,费马提出一个论断:27 是惟一的立方数,它比一个平方数大 2. 这也相当于说他猜想方程 $y^3=x^2+2$ 只有一组正整数解 $(x,y)=(5,3)$. 1770 年,欧拉给出这个问题的一种解法. 方程右边无法在整数范围内化成乘积形式,欧拉则引进复数 $\sqrt{-2}$,而把方程写成

$$y^3=x^2+2=(x-\sqrt{-2})(x+\sqrt{-2}).$$

于是欧拉考虑形如 $a+b\sqrt{-2}$ 的数,其中 a 和 b 是通常整数. 所有这种数组成的集合表示成 $\mathbf{Z}[\sqrt{-2}]$,易知这个集合和通常整数集合 $\mathbf{Z}$ 一样有加减乘运算. 在 $\mathbf{Z}[\sqrt{-2}]$ 中也可同样定义整除概念. 由 $y^3=x^2+2$ 易知 x 必为奇数,由此欧拉推出 $x-\sqrt{-2}$ 和 $x+\sqrt{-2}$ 是互素的. 但是它们的乘积是立方数 y^3,欧拉说,$x\pm\sqrt{-2}$ 一定是 $\mathbf{Z}[\sqrt{-2}]$ 中元素的立方,即

$$\begin{aligned}x+\sqrt{-2}&=(a+b\sqrt{-2})^3\\&=(a^3-6ab^2)+(3a^2b-2b^3)\sqrt{-2},\end{aligned}$$

其中 $a,b\in\mathbf{Z}$. 于是

$$x=a^3-6ab^2=a(a^2-6b^2),$$
$$1=3a^2b-2b^3=b(3a^2-2b^2).$$

由后一等式可知 $b=\pm1$,于是 $1=\pm(3a^2-2)$. 这又得出 $a=\pm1$. 因此 $x=(\pm1)\cdot(1-6)=\pm5$,而 $y=3$.

欧拉的这个解法是不严格的. 首先需要弄清为什么 $x-\sqrt{-2}$ 和 $x+\sqrt{-2}$ 互素(其中 x 是奇数)? 其次是:它们乘积为立方的时候,为什么每个都是立方? 这些事情是在 19 世纪初才由高斯解释清楚的. 但是欧拉把 $\mathbf{Z}$ 扩大成 $\mathbf{Z}[\sqrt{-2}]$ 来解不定方程的想法是非

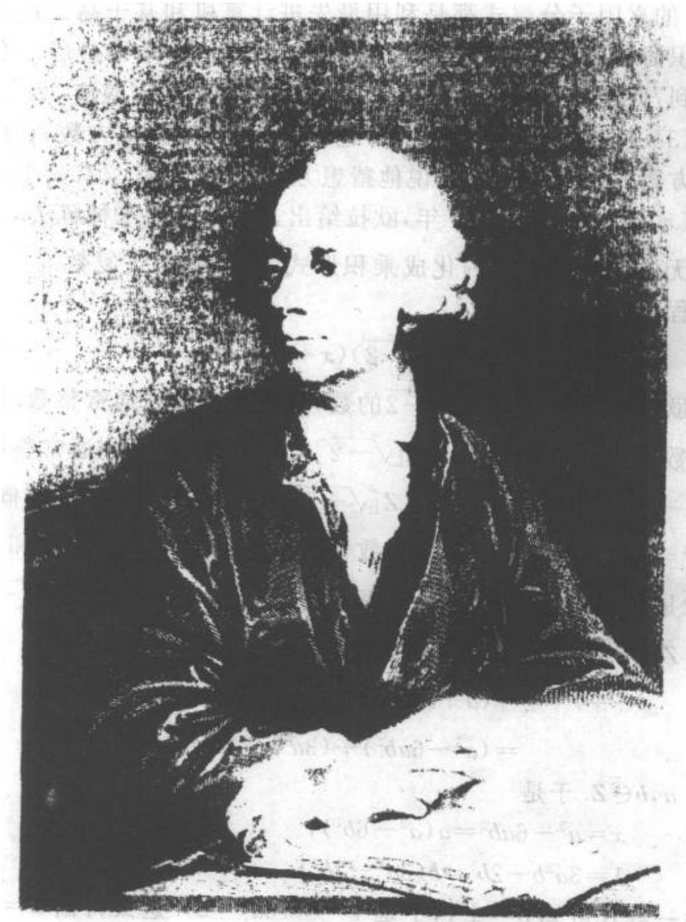


图 3　欧拉

常重要的. 在研究整数解问题时采用$\sqrt{-2}$的作法在很长时期不被欧拉同时代的人所接受，但这种新思想的萌芽后来发展成数论的一个重要分支.

（4）1640 年，费马在给 Mersenne 的信中断言：每个素数 $p=$

$4n+1$(即被 4 除余 1)都是二平方和(即两个平方数之和). 换句话说,当素数 $p\equiv1 \pmod 4$ 时,费马断言方程 $x^2+y^2=p$ 必有整数解. 1654 年,在给 Pascal 的信中费马又断言:奇素数 p 可以表成 x^2+2y^2 当且仅当 $p\equiv1,3 \pmod 8$,$p(p\neq3)$ 可以表成 x^2+3y^2 当且仅当 $p\equiv1 \pmod 3$. 费马声称他证明了这些论断,并且使用了无穷下降法. 事实上这些论断是后来于 1756 年由欧拉证明的,并且确实采用了无穷下降法,但是需要技巧.

接下来费马研究 x^2+5y^2. 在这里费马发生了困惑. 通过试验性的运算他意识到,素数 p 是否可写成 x^2+5y^2,应当和 p 模 20 的剩余有关,但是他没有提出像前面那样的准确性论断. 进而,他发现当 $p\equiv3,7 \pmod{20}$ 时,p 不能写成 x^2+5y^2 的形式,但是两个这样素数的乘积却可写成 x^2+5y^2 的形式,例如 $21=3\cdot7=1^2+5\cdot2^2$. 这种现象在过去没有发生过,例如对 x^2+y^2,当 $p\equiv3 \pmod 4$ 时 p 不是二平方和,并且这样的任意两个不同素数的乘积(例如 $21=3\cdot7$)也不是二平方和. 所以费马只提出如下的论断:若 p 和 q 是模 20 同余于 3 或 7 的两个素数,则 pq 可写成 x^2+5y^2 的形式.

欧拉没能证明这个猜想,因为他也不明白其中的原因. 但是欧拉在 1744 年提出了猜想:

设 p 为奇素数,则 p 可表成 x^2+5y^2 当且仅当 $p\equiv1,9 \pmod{20}$. 而 $2p$ 可表成 x^2+5y^2 当且仅当 $p\equiv3,7 \pmod{20}$.

这些猜想和其中的秘密是于 1773 年由拉格朗日证明和揭示的. 他开始考虑一般的二元二次型

$$ax^2+bxy+cy^2 \quad (a,b,c\in\mathbf{Z},\text{并且}(a,b,c)=1).$$

希望能够把这无穷多个二元二次型加以分类,使得每类中的二元二次型都可表示同样的整数. 容易想到的一种变换是变量之间的齐次线性变换:

$$\begin{cases}x'=\alpha x+\beta y,\\ y'=\gamma x+\delta y\end{cases} \quad (\alpha,\beta,\gamma,\delta\in\mathbf{Z}).$$

它把 $ax^2+bxy+cy^2$ 变成二元二次型 $a'x'^2+b'x'y'+c'y'^2$.

如果要求变换是可逆的，即

$$\begin{cases} x=\alpha'x'+\beta'y', \\ y=\gamma'x'+\delta'y' \end{cases} \quad (\alpha',\beta',\gamma',\delta'\in\mathbf{Z}).$$



图 4　拉格朗日

需要 $\alpha\delta-\beta\gamma=\pm1$. 满足这个条件的全部齐次线性变换形成群. 两个二元二次型可以由这样变换从一个变到另一个，称这两个二元二次型是等价的. 等价的二元二次型可以表示同样的整数. 即若两个二元二次型 $f(x,y)$ 和 $f'(x,y)$ 等价，则对每个整数 n，方程 $f(x,y)=n$ 有整数解当且仅当 $f'(x,y)=n$ 有整数解，换句话说，二元二次型 $f(x,y)$ 可以表示的整数全体，是 $f(x,y)$ 的等价类的不变量. 另一个等价类不变量为 $ax^2+bxy+cy^2$ 的判别式 b^2-4ac.

即若它与 $a'x^2+b'xy+c'y^2$ 等价，则 $b^2-4ac=b'^2-4a'c'$. 所有这些结论都可用初等的计算来验证（用群论和矩阵运算则更容易，不过拉格朗日时代还没有群论和矩阵论）.

综上所述，为了研究费马和欧拉提出的猜想，即二元二次型能表示哪些整数. 我们只需考虑它们的等价类. 由于等价的二元二次型有同样的判别式. 自然提出如下问题：

对每个整数 D，判别式为 D 的二元二次型有多少等价类，能否求出每个等价类的代表？

拉格朗日对于 $D<0$ 的情形很好地解决了这个问题. 他发现：在每个等价类中都有一个特别的二元二次型 $ax^2+bxy+cy^2$ 满足 $0\leqslant b\leqslant a\leqslant c$. 这样的二元二次型叫作约化的(reduced). 不同的约化型彼此不等价. 对于约化型我们有

$$-D=4ac-b^2\geqslant 4a^2-a^2=3a^2.$$

由于负整数 D 是固定的，所以满足比不等式的整数 a 只有有限多个，于是由 $0\leqslant b\leqslant a$ 知 b 也只有有限多个，而 c 由 $b^2-4ac=D$ 所决定. 这表明：判别式为 $D(<0)$ 的约化二元二次型只有有限多个. 所以判别式为 $D(<0)$ 的二元二次型只有有限多（等价）类. 这个类数表示成 $h(D)$，它等于满足

$$0\leqslant b\leqslant a\leqslant c,\quad b^2-4ac=D,\quad (a,b,c)=1$$

的整数 (a,b,c) 的组数.

由此容易算出，$h(-4)=h(-8)=h(-12)=1$，并且判别式为 $-4,-8,-12$ 的约化二元二次型分别为 x^2+y^2，x^2+2y^2 和 x^2+3y^2. 当 $D=-20$ 时，可求出判别式为 -20 的两个约化型 x^2+5y^2 和 $2x^2+2xy+3y^2$. 即 $h(-20)=2$. 这就是 x^2+5y^2 使费马受到困惑的特别之处：它有一个不等价的伴随者.

拉朗格日证明了欧拉的一个猜想：对于奇素数 p，p 可表示成 x^2+5y^2，当且仅当 $p\equiv 1,9 \pmod{20}$. 与此同时也证明了：p 可表示成 $2x^2+2xy+3y^2$ 当且仅当 $p\equiv 3,7 \pmod{20}$. 为了证明费马和欧拉的其他猜想，拉格朗日考虑二元二次型的合成. 首先我们有：

$$(x^2+5y^2)(x'^2+5y'^2)=(X^2+5Y^2),$$

其中 $X=xx'-5yy'$，$Y=xy'+x'y$. 这可以直接验证，但拉格朗日建议像欧拉做过的那样，采用$\sqrt{-5}$，则 $x^2+5y^2=(x-\sqrt{-5}y)(x+\sqrt{-5}y)=|x+\sqrt{-5}y|^2$，于是

$$\begin{aligned}(x^2+5y^2)(x'^2+5y'^2)&=|x+\sqrt{-5}y|^2\cdot|x'+\sqrt{-5}y'|^2\\&=|(x+\sqrt{-5}y)(x'+\sqrt{-5}y')|^2\\&=|(xx'-5yy')+\sqrt{-5}(xy'+x'y)|^2\\&=(xx'-5yy')^2+5(xy'+x'y)^2.\end{aligned}$$

由此可知，若整数 n 和 m 均可表成 x^2+5y^2，则它们的乘积 nm 也可如此. 拉格朗日把上式左边看作是两个二元二次型 x^2+5y^2 的"合成"，结果仍是 x^2+5y^2. 同样地，他给出：x^2+5y^2 和 $2x^2+2xy+3y^2$ 的合成为 $2x^2+2xy+3y^2$，而两个 $2x^2+2xy+3y^2$ 的合成为 x^2+5y^2：

$$(x^2+5y^2)(2x'^2+2x'y'+3y'^2)=2X^2+2XY+3Y^2,$$

其中 $X=xx'+xy'-yx'-3yy'$，$Y=xy'+2yx'+yy'$.

$$(2x^2+2xy+3y^2)(2x'^2+2x'y'+3y'^2)=X^2+5Y^2,$$

其中 $X=2xx'+xy'+yx'-2yy'$，$Y=xy'+yx'+yy'$.

拉格朗日实际给出判别式为 -20 的两个约化二元二次型 $A=x^2+5y^2$ 和 $B=2x^2+2xy+3y^2$ 之间的一种合成运算：$AA=BB=A$，$AB=BA=B$. 后来他于 1798 年又一般性地证明了：判别式为负整数 D 的任意两个二元二次型仍可合成为判别式为 D 的二元二次型. 到了 19 世纪，高斯进一步证明判别式为 D 的二元二次型等价类对于合成运算形成有限（阶数为 $h(D)$）交换群.

拉格朗日利用二元二次型 $A=x^2+5y^2$ 和 $B=2x^2+2xy+3y^2$ 的合成结果很容易证明费马和欧拉的猜想：由于拉格朗日证明了 $p=20n+3$ 和 $20n+7$ 可用 B 表示，从而这样两个素数的乘积可用 $B\cdot B=A$ 表示，这就是费马的猜想. 另一方面，2 显然可由 B 表示，所以 $2p$ 也可由 $B\cdot B=A$ 表示，这就是欧拉的猜想.

除了费马和拉格朗日，18世纪另一个重要的法国数论学家是勒让得. 他发现了二次互反律，这是后来数论发展的一个重要源头. 他对于方程 $ax^2+by^2+cz^2=0$（其中 a,b,c 是非零整数），得到有非零整数解的充分必要条件，这是近代代数数论中关于二次型方程局部-整体原则的一个重要的例子. 他提出算术级数中素数的如下猜想：设 m 和 l 是互素的正整数，则存在无穷多个素数 $p\equiv l(\mathrm{mod}\ m)$. 这个猜想到了19世纪才由德国人狄里赫利所证明，并引发出研究数论的解析方法.



图5 勒让得 Adrien Marie Legendre

17世纪费马对于数论提出了一系列猜想，促使18世纪的欧拉、拉格朗日和勒让得对数论的研究，欧拉对费马的工作研究了40年之久. 这些研究孕育着数论在19世纪将会有一个大的发展. 关于上述诸位数论学家的生平和工作的更详细介绍可参见下列文献.

参 考 文 献

[1] W. Scharlau and H. Opolka, From Fermat to Minkowski:

Lectures on the Theory of Numbers and It's Historical Development. Springer-Verlag, 1984.

[2] A. Weil, Number Theory, An approach through history. From Nammurapi to Legendre. Birkhauser, 1983.

第二章　经典代数数论：从高斯到希尔伯特

§2.1　高斯和他的《数论探究》(1801年)

19世纪数论取得重大的进步，数论从少数人的职业和业余爱好变成数学的一个重要和主流学科. 数论不仅在理论上系统化，而且研究手段有巨大的创新，产生了数论两个新的重要分支：代数数论和解析数论. 19世纪数论的发展和代数数论的建立都是从高斯在24岁时出版的《数论探究》一书开始的.

高斯(Gauss，1777—1855)在这本书的前言中说：欧几里德和丢番都研究了数论的许多特殊规律，但是数论需要有"一般原则". 这些一般原则已经由费马、欧拉、拉格朗日和勒让德开始研究. 他在这本书中要系统介绍自己的数论研究结果.

《数论探究》一书共分七章. 前四章分别是同余、一次同余、幂剩余和原根指数、二次同余式和二次互反律. 在这一部分，高斯采用了沿用至今的同余式符号，系统讲述了同余式运算法则和模 m 同余类的概念，严格证明了一系列重要的同余定理(费马小定理、欧拉定理等)，系统论述了原根和指数的概念和规律. 用现在的语言，就是研究了有限阿贝尔加法群 $\mathbf{Z}/m\mathbf{Z}$ 和乘法群 $(\mathbf{Z}/m\mathbf{Z})^{\times}$ 的结构. 他严格证明了 $(\mathbf{Z}/p\mathbf{Z})^{\times}$ 是乘法循环群(其中 p 为素数)，生成元即是模 p 的原根. 原根的存在性欧拉就已经知道，而高斯利用指

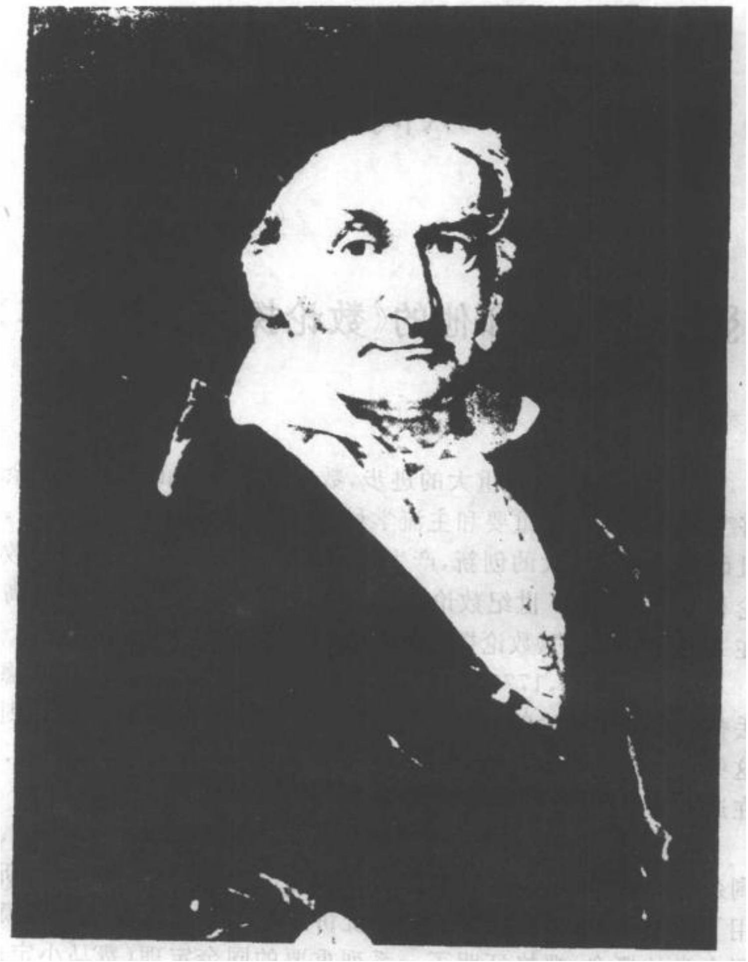


图 6　高斯(Carl Friedrich Gauss)

数理论计算出模 p 的原根(即 $p-1$ 阶乘法循环群 $(Z/pZ)^{\times}$ 的生成元)共有 $\varphi(p-1)$ 个. 在这一部分高斯还系统论述了一次和二次同余方程的解法,并且用来求一次和二次不定方程的整数解. 高斯

叙述了一次不定方程求整数解的完整理论.对于一元二次同余,高斯采用配方法把问题归结为最基本情形 $x^2\equiv a(\mathrm{mod}\ p)$. 由此建立了二次剩余概念和系统理论,其核心部分是证明了二次互反律.二次互反律的建立和高斯以及后人对它的深刻思考,是代数数论乃至现代数论的一个重要源头,我们在后面将陆续介绍.

第五和第六章论述二次不定方程 $ax^2+bxy+cy^2=n$ 的整数解问题和各种具体应用(其中 $a,b,c,n\in\mathbf{Z}$). 这部分占了该书的2/3篇幅.如果上面二次不定方程有整数解,便称二元二次型 $f(x,y)=ax^2+bxy+cy^2$ 表示整数 n.

高斯研究下列一般性问题:一个固定的二元二次型 $f(x,y)$可以表示哪些整数?

高斯采用了更宏观的视野,即不是对二元二次型逐个加以研究,而是对所有二元二次型加以分类.他沿用第一章介绍的拉格朗日方法,用行列式为±1 的二阶整数方阵群的作用,把二元二次型分成等价类.判别式 $\Delta=b^2-4ac$ 是等价不变量,并且等价的二元二次型可以表示同样的整数.进而,高斯发展了拉格朗日的约化方法,证明了对每个非零整数 Δ,判别式为 Δ 的二元二次型只有有限个等价类.当 $\Delta<0$(即 $ax^2+bxy+cy^2$ 是正定二次型($a,c>0$))时,证明是容易的并且本质上由拉格朗日给出;当 $\Delta>0$(即 $ax^2+bxy+cy^2$ 是不定型)时,高斯修改了约化二元二次型的定义,证明克服了很多困难.以 $h(\Delta)$表示判别式为 Δ 的二元二次型等价类数.高斯对许多 Δ 计算了 $h(\Delta)$的值.例如 $h(-20)=2$,这相当于拉格朗日的发现,即 x^2+5y^2 和 $2x^2+2xy+3y^2$ 是不等价的.高斯还计算出:当 $d=1,2,3,7,11,19,43,67$ 和 163 时,$h(-d)=1$,并且猜想对其它负整数 Δ,均有 $h(\Delta)>1$.

这个猜想在 1967 年才由英国数学家 Baker 和美国数学家 Stark 各自独立地加以证明(Baker 使用超越方法,Stark 采用模形式理论).

另一方面,高斯还猜想:存在无限多个正整数 Δ,使得

$h(\Delta)=1$. 这个猜想至今未解决.

高斯还把拉格朗日关于二元二次型的合成运算定义作了修改,证明这是二元二次型等价类的运算,并且表明这个运算满足结合律和交换律. 用现在的语言来说,高斯论述了:所有判别式为 Δ 的二元二次型等价类似于合成运算形成 $h(\Delta)$阶阿贝尔群. 高斯还给出计算 $h(\Delta)$的解析公式. 最后,高斯还建立了二元二次型的种(genus)理论. 按照这种理论可以证明,如果 Δ 有 r 个不同的素因子,则 $2^{r-1}|h(\Delta)$. 所有上述结果都是十分深刻的理论成果. 到了19 世纪后半期,狄德金(Dedekind)建立理想论之后,高斯这些结果直接转化成二次域理想分解和理论类群的深刻结果,成为经典代数数论的重要内容. 按照高斯关于二元二次型合成运算的定义,为了验证运算满足结合律,需要核准 27 个等式的正确性,高斯把其中大部分等式的核准工作留给读者. 在狄德金采用理想论的语言之后,运算的结合律和交换律则变成了显而易见的事情.

《数论探究》的前六章讲述了纯粹数论内容,而第七章则研究一个几何问题:对哪些正整数 $n\geqslant 3$,可以用圆规直尺作出一个正 n 边形来?这也相当于问:对哪些 n 可以用圆规直尺把单位圆周 n 等份?在复平面上把单位圆周 n 等份,如果圆心在原点而一个分点为 1,则 n 个分点为 $S_n^i(i=0,1,2,\cdots,n-1)$,其中

$$S_n=e^{\frac{2\pi i}{n}}=\cos\frac{2\pi}{n}+i\sin\frac{2\pi}{n}.$$

古希腊人就知道如何用圆规直尺作出正三角形,正方形和正五边形,而且知道如何用圆规直尺作角的平分线,所以若用尺规可作正 n 边形,则也可作出正 $2n$ 边形. 高斯在他 19 岁的时候,出色地给出了用尺规画出正 17 边形的方法. 在书的第七章高斯证明了:若 $n=2^m p_1 p_2\cdots p_t$,其中 $m\geqslant 0$,而 p_i 均是形如 $2^{2^l}+1$ 的不同素数(即费马素数),则用尺规可以构作出正 n 边形.

到了 19 世纪中期产生了伽罗华理论之后,人们知道高斯已经找到了可用尺规作出的全部正 n 边形.

让我们看一下高斯是如何证明上述结果的. 首先，利用初等数论不难证明：若用尺规可作出正 n 边形和正 m 边形，并且 m 和 n 互素，则用尺规可构作正 mn 边形. 于是问题归结为要证：对每个费马素数 $p=2^{2^l}+1$，用尺规均可构作正 p 边形. 为此，高斯研究

$$g=\sum_{a=1}^{p-1}\left(\frac{a}{p}\right)\zeta_p^a \quad \left(\zeta_p=e^{\frac{2\pi i}{p}}\right),$$

其中 $\left(\frac{a}{p}\right)$ 为勒让得符号，即对于与 p 互素的整数 a，

$$\left(\frac{a}{p}\right)=\begin{cases}1, & \text{若 } a \text{ 为模 } p \text{ 的二次剩余,}\\ -1, & \text{若 } a \text{ 为模 } p \text{ 的二次非剩余.}\end{cases}$$

g 现在叫作**高斯和**(在此之前拉格朗日研究过与高斯和类似的数，叫拉格朗日结式). 高斯利用勒让得符号关系 $\left(\frac{a}{p}\right)\left(\frac{b}{p}\right)=\left(\frac{ab}{p}\right)$ 证明了

$$g^2=\left(\frac{-1}{p}\right)p=\begin{cases}p, & \text{若 } p\equiv 1 \pmod 4,\\ -p, & \text{若 } p\equiv 3 \pmod 4.\end{cases}$$

于是 $g=\pm\sqrt{\left(\frac{-1}{p}\right)p}$. 进而高斯通过不平凡的努力决定出此式右边的符号恒为正，即 $g=\sqrt{p}$ 或 $\sqrt{p}\,i$. 如果令

$$A_1=\sum_{\substack{1\leqslant a\leqslant p-1\\ \left(\frac{a}{p}\right)=1}}\zeta_p^a,\quad A_2=\sum_{\substack{1\leqslant a\leqslant p-1\\ \left(\frac{a}{p}\right)=-1}}\zeta_p^a,$$

则 $A_1+A_2=\sum\limits_{1\leqslant a\leqslant p-1}\zeta_p^a=-1$，$A_1-A_2=g=\sqrt{p}$ 或 $\sqrt{p}\,i$. 于是

$$A_1=\frac{1}{2}(\sqrt{p}-1) \quad \text{或} \quad \frac{1}{2}(\sqrt{p}\,i-1),$$

$$A_2=\frac{1}{2}(-\sqrt{p}-1) \quad \text{或} \quad \frac{1}{2}(-\sqrt{p}\,i-1).$$

由平面几何知道，如果确定一个单位长度(即长度为 1)的线段，那么用尺规可作出长度为任何正有理数的线段，即可作出任何正有

理数.更一般地,如果用尺规可作出复数 α 和 β(即复平面上的点 α 和 β),则可作出 α 和 β 的四则运算数 $\alpha\pm\beta,\alpha\beta$ 和 $\alpha/\beta(\beta\neq 0)$以及开平方 $\sqrt{\alpha}$(事实上,尺规作图的能力也仅限于此).于是我们便知道,点 A_1 和 A_2 是可用尺规作出的.现在我们设 $p=2^{2^l}+1$ 是费马素数,记 $2^l=m$.取 g 为模 p 的一个原根,则模 p 的$\frac{p-1}{2}$个二次剩余为 $g^{2i}\left(i=0,1,\cdots,\frac{p-3}{2}\right)$,而$\frac{p-1}{2}$个非二次剩余为 g^{2i+1} $\left(i=0,1,\cdots,\frac{p-3}{2}\right)$.所以 A_1 和 A_2 均是$\frac{p-1}{2}=2^{m-1}$个 p 次单位根之和:

$$A_1=\sum_{i=0}^{\frac{p-3}{2}}\zeta_p^{g^{2i}},\ A_2=\sum_{i=0}^{\frac{p-3}{2}}\zeta_p^{g^{2i+1}},$$

进而,模 p 的四次剩余共有$\frac{p-1}{4}=2^{m-2}$个:$g^{4i}(i=0,1,\cdots,2^{m-2}-1)$.

令

$$B_1=\sum_{i=0}^{2^{m-2}-1}\zeta_p^{g^{4i}},\ B_2=A_1-B_1,$$

则 B_1 和 B_2 均是 2^{m-2}个 p 次单位根之和.并且 $B_1+B_2=A_1$.高斯证明了 $B_1B_2=A'$可以用有理数以及 A_1 和 A_2 通过四则运算表达出来.从而 $B_1+B_2=A_1$ 和 $B_1B_2=A'$都可由尺规作出.而 B_1 和 B_2 是二次方程 $x^2-A_1x+A'=0$ 的解,由求解公式中只有四则运算和开平方运算.所以 B_1(和 B_2)可由尺规作出.可同样证明 2^{m-3}个 p 次单位根之和

$$C=\sum_{i=0}^{2^{m-3}-1}\zeta_p^{g^{8i}}=\zeta_p+\zeta_p^{g^8}+\cdots$$

也可由尺规作出.继续下去,最后便得到 ζ_p 可由尺规作出,于是用尺规可作出正 p 边形,其中 p 为费马素数.这就是高斯的证明.

高斯和以及它的各种推广，成为后来数论研究的重要工具. 在 18 世纪，欧拉研究费马猜想 $n=3$ 的情形时，使用了三次单位根 $\zeta_3=\frac{1-\sqrt{3}i}{2}$. 而高斯则更进一步研究了单位根 ζ_p 的性质，可以看成是研究分圆域 $\boldsymbol{Q}(\zeta_p)$ 的开端. 比如说，由等式 $g=\sqrt{p}$ 或 $\sqrt{-p}$ 可知二次域 $\boldsymbol{Q}(\sqrt{\pm p})$ 是分圆域 $\boldsymbol{Q}(\zeta_p)$ 的子域（因为 g 可由 p 次单位根表达出来），从而不难推出任何二次域都是分圆域的子域，这就引发出一个至今未解决的一个代数数论问题（希尔伯特第 12 问题）. 这表明，虽然第七章研究的是几何问题，但它实质上讲述了代数数论（分圆域）的内容. 分圆域的更深入研究是由库默尔在 19 世纪中期作出的（见下一节）.

以上是《数论探究》一书的主要内容. 为了进一步说明它对代数数论的建立所起的作用，我们再结合高斯（以及其他人）后来的工作，对其中两个方面作专门的解释，这两个方面是高斯数域 $\boldsymbol{Q}(i)$ 和二次互反律.

在高斯之前，欧拉和拉格朗日、勒让德等人在研究 Pell 方程 $x^2-dy^2=1$ 和费马方程 $x^3+y^3=z^3$ 时采用了环 $\boldsymbol{Z}[\sqrt{d}]$ 和 $\boldsymbol{Z}[\zeta_3]$. 特别是欧拉在证明 $x^3+y^3=z^3$ 没有正整数解时事实上使用了环 $\boldsymbol{Z}[\zeta_3]$ 的惟一因子分解特性. 但是对二次域 $\boldsymbol{Q}(\sqrt{d})$ 的系统研究特别是对环 $\boldsymbol{Z}[\sqrt{d}]$ 的惟一因子分解性质加以认真研究的是高斯. 这种研究是在 1837 年进行的. 高斯在《数论探究》中用二元二次型分类的语言讨论一般二次方程 $ax^2+bxy+cy^2=n$ 的整数解. 但是他于 1837 年讨论特殊方程 $x^2+y^2=n$ 时，则深入研究和使用了环 $\boldsymbol{Z}[i]$ 的惟一因子分解性质. 这个环现在被称为**高斯整数环**，其中每个元素 $a+bi\,(a,b\in\boldsymbol{Z})$ 叫**高斯整数**.

类似于整数环 $\boldsymbol{Z}$ 的情形，高斯定义了高斯整数之间的整除概念：高斯整数 α 可整除 β，是指存在高斯整数 γ，使得 $\alpha\gamma=\beta$（表示成 $\alpha\mid\beta$）. $\boldsymbol{Z}[i]$ 中乘法可逆元素（或叫作**单位**）共有 4 个：±1，$\pm i$. 定

义 $\alpha=a+bi$ 的范数为 $N(\alpha)=|a+bi|^2=a^2+b^2$，则高斯整数 α 为单位当且仅当 $N(\alpha)=1$. 由于方程 $n=x^2+y^2$ 相当于 $n=(x+iy)(x-iy)=N(x\pm iy)$，可知 n 为二平方和当且仅当 n 在环 $\mathbf{Z}[i]$中分解成两个(共轭)高斯整数之积. 这就引发出高斯整数环 $\mathbf{Z}[i]$中的因子分解问题. 而本质上不能分解的高斯整数 π 被称作高斯素数. 确切地说：π 是高斯素数，是指 π 不是 0 或单位，并且若 $\pi=\alpha\beta$，则高斯整数 α 和 β 当中必有一个为单位.

设 α 和 β 是两个非零高斯整数. 如果 $\alpha|\beta$ 并且 $\beta|\alpha$，则称 α 和 β 等价，表示成 $\alpha\sim\beta$. 易知这相当于 $\alpha=\beta\gamma$，其中 γ 为单位. 于是和 α 等价的恰有四个高斯整数 $\pm\alpha$ 和 $\pm\alpha i$. 现在可以叙述高斯的重要结果.

高斯整数环是惟一因子分解整环(简记为 UFD，Uniquely Factorization Domain). 也就是说：设 $\alpha\in\mathbf{Z}[i]$，$\alpha\neq 0,\pm1,\pm i$，则

(1) (存在性)α 可以表示成有限个高斯素数的乘积；

(2) (惟一性)如果 $\alpha=\pi_1\cdots\pi_r=\pi'_1\cdots\pi'_s$，其中 π_i 和 π'_j 均是高斯素数，则 $r=s$，并且适当调换诸 $\pi'_j(1\leqslant j\leqslant r)$的次序，可以使 $\pi_i\sim\pi'_i(1\leqslant i\leqslant r)$.

整数环 $\mathbf{Z}$ 的惟一因子分解特性是由欧几里德基本上建立而由高斯严格证明的. 对于 $\mathbf{Z}[i]$高斯采用了类似的证明方法. 存在性的证明是容易的：如果 α 是高斯素数则分解完毕. 否则 $\alpha=\beta\gamma$，其中 $\beta,\gamma\in\mathbf{Z}[i]$并且均不为单位. 于是 $N(\beta)$和 $N(\gamma)$均不为 1，但是 $N(\alpha)=N(\beta)\cdot N(\gamma)$，因此 $N(\beta)$和 $N(\gamma)$都是比 $N(\alpha)$小的正整数. 这样下去便知有限步之后，α 必可分解为有限个高斯素数的乘积. 为了证明分解的惟一性，需要像 $\mathbf{Z}$ 中那样证明素数的如下性质：若 $\pi,\alpha,\beta\in\mathbf{Z}[i]$并且 π 是高斯素数. 如果 $\pi|\alpha\beta$，则 $\pi|\alpha$ 或者 $\pi|\beta$. 而这件事可以由下面的结果推出：

(高斯整数环的欧几里德除法算式)　设 $\alpha,\beta\in\mathbf{Z}[i]$，$\beta\neq0$，则存在惟一的 $q,\gamma\in\mathbf{Z}[i]$，使得

$$\alpha=q\beta+\gamma, 0\leqslant N(\gamma)<N(\beta).$$

为了证明这件事,高斯使用了$\mathbf{Z}[i]$所在的二次域$\boldsymbol{Q}(i)$(叫作高斯数域). 令$\frac{\alpha}{\beta}=a+bi$,其中$a,b\in\boldsymbol{Q}$(有理数域).

令$q=(a)+(b)i$,其中(a)和(b)分别是距离a和b最近的整数,则$q\in\mathbf{Z}[i]$,而$N\left(\frac{\alpha}{\beta}-q\right)\leqslant\left(\frac{1}{2}\right)^2+\left(\frac{1}{2}\right)^2<1$. 因此对$\gamma=\alpha-q\beta\in\mathbf{Z}[i]$,便有$N(\gamma)=N(\beta)\cdot N\left(\frac{\alpha}{\beta}-q\right)<N(\beta)$.

既然高斯整数惟一分解为高斯素数之积,下一个问题自然需要决定出所有的高斯素数. 首先,如果$\alpha\in\mathbf{Z}[i]$,$N(\alpha)=p$是通常的素数,则α必是高斯素数(若$\alpha=\beta\gamma$,则$p=N(\alpha)=N(\beta)\cdot N(\gamma)$,于是$N(\beta)$和$N(\gamma)$必有一个为1,即$\beta$和$\gamma$必有一个为单位). 其次,每个通常的奇素数$p$在比$\mathbf{Z}$更大的环$\mathbf{Z}[i]$中可能不是高斯素数. 如果$p=\alpha\beta$,其中$\alpha,\beta\in\mathbf{Z}[i]$,则$p^2=N(p)=N(\alpha)\cdot N(\beta)$. 如果$\alpha$和$\beta$不是单位,则必然$N(\alpha)=N(\beta)=p$. 于是$\alpha$和$\beta$均是高斯素数,易知$\alpha=a+bi$而$\beta=\bar{\alpha}=a-bi$. 于是$p=a^2+b^2$,即$p$为二平方和. 由欧拉结果这相当于$p\equiv1\pmod 4$. 这就表明:若$p\equiv1\pmod 4$,则$p$为两个高斯素数之积$p=\pi\bar{\pi}$,并且易知$\pi$和$\bar{\pi}$不等价. 如果$p\equiv3\pmod 4$,则$p$仍是高斯素数. 最后$2=(1+i)\cdot(1-i)$,其中$1+i$和$1-i$是等价的高斯素数$\left(\text{因为}\frac{1+i}{1-i}=i\text{为单位}\right)$. 以上方式便给出全体高斯素数.

利用$\mathbf{Z}[i]$的惟一因子分解特性和p在$\mathbf{Z}[i]$中的分解情况,高斯得到二平方和问题的最完整结果. 对每个正整数n都可惟一表达成$n=n_1^2\cdot n_2$,其中n_2是1或者是不同素数之积. n_2叫n的无平方因子部分.

关于二平方和的高斯定理:设n_2是正整数n的无平方因子部

分. 则 n 是二平方和当且仅当 n_2 没有素因子 $p\equiv 3 \pmod 4$.

利用 $\mathbf{Z}[i]$ 的惟一因子分解特性，高斯还给出了方程 $x^2+y^2=n$ 整数解个数的一般公式. 类似地可以证明环 $\mathbf{Z}[\zeta_3]$ 也是 UFD，从而对欧拉证明 $x^3+y^3=z^3$ 没有正整数解最终给出了严格的基础. 后来，高斯研究了一般环 $\mathbf{Z}[\sqrt{d}]$ 中的因子分解特性，发现有许多环的因子分解不满足惟一性. 例如环 $\mathbf{Z}[\sqrt{-5}]$，$6=2\cdot 3=(1+\sqrt{-5})(1-\sqrt{-5})$，其中 2，3 和 $1\pm\sqrt{-5}$ 都是环 $\mathbf{Z}[\sqrt{-5}]$ 中的"素数"，但是 2 和 $1\pm\sqrt{-5}$ 不等价$\left(\text{因为}\frac{1\pm\sqrt{-5}}{2}\right.$都不是此环中的元素). 由此引发出高斯对二次域的一系列深刻的研究，成为代数数论的重要组成部分.

最后谈谈二次互反律. 高斯在《数论探究》第 107 节中说："给了数 m，容易判别哪些数是模 m 的二次剩余和二次非剩余. 但反问题要难得多：对给定的 m，试问对哪些奇素数 p，m 是模 p 的二次剩余？"

费马在 1642 年就知道：对每个奇素数 p，$\left(\frac{-1}{p}\right)=1$ 当且仅当 $p\equiv 1 \pmod 4$，这是由欧拉于 1760 年证明的，费马还猜想：$\left(\frac{2}{p}\right)=1$ 当且仅当 $p\equiv\pm 1 \pmod 8$，勒让得于 1775 年给出了第一个证明. 高斯在 1796 年 19 岁的时候第一次证明了如下漂亮的结果（它是由勒让德于 1785 年提出来的）：

二次互反律：设 p 和 q 为不同的奇素数，则

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right)=(-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}.$$

换句话说，若 $p\equiv q\equiv 3 \pmod 4$，则 $\left(\frac{p}{q}\right)=-\left(\frac{q}{p}\right)$. 否则 $\left(\frac{p}{q}\right)=$

$\left(\frac{q}{p}\right)$.

高斯非常喜欢他证明的这个二次互反律. 在《数论探究》一书中给出了六个证明,后来又给出了两个证明(近来有人统计二次互反律的证明已超过一百个). 第一个证明采用了巧妙的初等方法,而第六和第八个证明利用了高斯和(即使用了单位根 ζ_p).

考虑

$$g_q=\sum_{a=1}^{p-1}\left(\frac{a}{p}\right)\zeta^{qa},$$

易知 $g_q=\left(\frac{q}{p}\right)g_1$,而前面已说过 $g_1^2=\left(\frac{-1}{p}\right)p$. 由于

$$\begin{aligned} g_1^q &= \left(\sum_{a=1}^{p-1}\left(\frac{a}{p}\right)\zeta^a\right)^q \equiv \sum_{a=1}^{p-1}\left(\frac{a}{p}\right)^q\zeta^{aq} \pmod q \\ &\equiv \sum_{a=1}^{p-1}\left(\frac{a}{p}\right)\zeta^{aq}=g_q=\left(\frac{q}{p}\right)g_1 \pmod q, \end{aligned}$$

而 $g_1^2=\left(\frac{-1}{p}\right)p$ 与 q 互素,可知 $g_1^{q-1}\equiv\left(\frac{q}{p}\right)\pmod q$.

但是 $\qquad g_1^{q-1}=\left(\frac{-1}{p}\right)^{\frac{q-1}{2}}p^{\frac{q-1}{2}}\equiv\left(\frac{-1}{p}\right)^{\frac{q-1}{2}}\left(\frac{p}{q}\right) \pmod q,$

(这里用了欧拉一个结果:$p^{\frac{q-1}{2}}\equiv\left(\frac{p}{q}\right)\pmod q$). 于是

$$\left(\frac{q}{p}\right)\equiv\left(\frac{-1}{p}\right)^{\frac{q-1}{2}}\left(\frac{p}{q}\right) \pmod q.$$

由于同余式两端为±1 而 q 为奇素数,从而必然$\left(\frac{q}{p}\right)=\left(\frac{-1}{p}\right)^{\frac{q-1}{2}}\cdot\left(\frac{p}{q}\right)$. 再由$\left(\frac{-1}{p}\right)=(-1)^{\frac{p-1}{2}}$即得二次互反律. 高斯的第六和第八证明本质上是上述证明的两种不同形式.

二次互反律是环 $\mathbf{Z}$ 上的一个论断,而且高斯也给出了初等(即在有理数域 $\mathbf{Q}$ 内)的证明. 人们对它不断地寻求新的证明方

法，而且使用超出有理数域范围的各种手段，主要有两个原因，一是这个结果非常漂亮和吸引人、二是希望对它有更深层次的理解和推广. 爱森斯坦(Eisenstein，1823—1852)利用三角函数给出二次互反律的一个证明，如果把三角函数改成某种椭圆函数，则用同样方法可证明三次和四次互反律. 为了确切叙述三次和四次互反律的结果，需要利用环 $\boldsymbol{Z}[\omega]$和 $\boldsymbol{Z}[i]$，其中 $\omega=\zeta_3$ 和 $i=\sqrt{-1}=\zeta_4$. 以三次互反律为例，设 $p\equiv1(\bmod 3)$，g 为模 p 的一个原根，则 g 是 $p-1=3m(m\in\boldsymbol{Z})$阶元素(对于有限域 $\boldsymbol{F}_p=\boldsymbol{Z}/p\boldsymbol{Z}$ 中的乘法)，并且对每个与 p 互素的整数 a，均有 $l(\bmod 3m)$，使得 $a\equiv g^l(\bmod p)$. 定义

$$\left(\frac{a}{p}\right)_3=\begin{cases}1, & \text{若 } l\equiv0(\bmod 3)\\ \omega, & \text{若 } l\equiv1(\bmod 3)\\ \omega^2, & \text{若 } l\equiv2(\bmod 3)\end{cases}$$

换句话说，$\left(\frac{a}{p}\right)_3=\omega^l$. 则 a 是模 p 的三次剩余，当且仅当$\left(\frac{a}{p}\right)_3=1$. 这件事是与原根 g 的选取方式无关的.

如果 a 不是模 p 的三次剩余，则$\left(\frac{a}{p}\right)_3$ 为 ω 还是 ω^2 依赖于原根 g 的选取方式. 我们无法在有理数域范围内说清楚 $\varphi(p-1)$个原根中哪一个是特别的，但是利用 $\boldsymbol{Z}[\omega]$的惟一因子分解性质可以选取模 p 一个特定的原根，并且证明三次互反律. 同样地，利用 $\boldsymbol{Z}[i]$的惟一因子分解性质可以确切叙述和证明四次互反律. 四次互反律的形式由高斯于 1832 年给出，但是三次和四次互反律的第一个严格证明是由爱森斯坦于 1844 年给出的(利用高斯和)，详见[IR82]一书第九章(爱森斯坦只活了 29 年，他在 1844 年 21 岁的时候，一年之内发表了 25 篇文章，其中包括三次和四次互反律的证明，以及我们在后面将要介绍的关于椭圆模函数和模形式的重要结果). 另一个方向则是希望把二次互反律从 $\boldsymbol{Z}$ 推广到更为一般的环上. 19 世纪中期，高斯的学生狄里赫利(Dirichlet，1805—

1859)给出高斯数域 $Q[i]$上的二次互反律,到了 19 世纪末期,希尔伯特给出了任意代数数域上的二次互反律,并且用分圆域中素理想分解法则给出高斯互反律一个证明,用以解释二次互反律相当于二次域中的素理想分解律. 于是希尔伯特问:对于任意代数数域应当有怎样的一般互反律?(希尔伯特第 9 问题)这个问题不仅推动了 20 世纪代数数论的发展,而且一直是近 30 年来现代数论一个最热门的研究课题(朗兰兹纲领).

综上所述我们可以看出,高斯在 19 世纪前期对数论所作的一系列研究工作不仅促使代数数论的产生和发展,而且对现代数论研究仍有深刻的影响力.

参 考 文 献

[1] C.-F. Gauss, Disquisitiones Arithmeticae, 1801(英文版 1966, Yale Univ, Press).

§2.2 库默尔研究费马猜想(1847 年)

继高斯之后,对代数数论的建立作了重大贡献的是另一个德国人库默尔(Kummer,1810—1893). 他在研究费马猜想时对于分圆域的理论得到了一系列重要结果.

1770 年欧拉用环 $\mathbf{Z}[\omega]$的知识证明了费马猜想(对每个奇素数 p, $x^p+y^p=z^p$ 没有正整数解)的 $p=3$ 情形. 到了 19 世纪,勒让德和狄里赫利分别于 1825 和 1828 年独立地证明了 $p=5$ 的情形,

Lamé 于 1839 年证明了 $p=7$ 的情形，证明于 1840 年由勒让德简化，但是证明仍旧十分复杂. 他们的证明原则上要使用环 $\mathbf{Z}[\zeta_5]$和 $\mathbf{Z}[\zeta_7]$的惟一因子分解性，而这在当时并未被严格证明. 1847 年，法国人 Lamé 在巴黎科学院的学术例会上宣布他证明了费马猜想. 而柯西(Cauchy)当即指出：Lamé 的想法他早就有所考虑，但是要用到环 $\mathbf{Z}[\zeta_p]$的惟一因子分解性质，并且柯西认为对于某些 p，$\mathbf{Z}[\zeta_p]$可能不是 UFD. 这场争论一直继续到三周之后库默尔的文章从德国寄到巴黎，便安静下来.

库默尔在文章中说，他早在 1844 年就证明了

(1) 如果 $\mathbf{Z}[\zeta_p]$是惟一因子分解整环(UFD)，则费马猜想对 p 成立.

(2) 当 $p\leqslant 19$ 时，$\mathbf{Z}[\zeta_p]$是 UFD，而 $\mathbf{Z}[\zeta_{23}]$不是 UFD. 换句话说，库默尔用统一的方法证明了：对所有奇素数 $p\leqslant 19$，费马猜想都正确.

让我们简要介绍一下库默尔的证明思想. 首先，他把方程在 $\mathbf{Z}[\zeta_p]$中化成乘积形式：

$$x^p=z^p-y^p=(z-y)(z-\zeta_p y)\cdots(z-\zeta_p^{p-1}y).$$

如果方程有正整数解(x,y,z)，不妨设 x,y,z 两两互素，他证明了方程右边 p 个因子在环 $\mathbf{Z}[\zeta_p]$中也是两两互素的. 如果 $\mathbf{Z}[\zeta_p]$是 UFD，由于方程左边 x^p 是 p 次方，所以每个因子 $z-\zeta_p^i y$ 也本质上是(即等价于)$\mathbf{Z}[\zeta_p]$中元素的 p 次方，即

$$z-\zeta_p^i y=\varepsilon_i\alpha_i^p\,(0\leqslant i\leqslant p-1),$$

其中 ε_i 是环 $\mathbf{Z}[\zeta_p]$中单位，而 $\alpha_i\in\mathbf{Z}[\zeta_p]$. 这虽然是很强的条件，但是要推导出矛盾(从而对 p 证明费马猜想)库默尔克服了一系列重大困难. 第一个困难是条件中出现单位 ε_i. $\mathbf{Z}[\zeta_p]$的单位(乘法)群 U_p 是秩为$\frac{p-3}{2}$的阿贝尔群. 当 $p\geqslant 5$ 时这是很大的无限群. 决定整个单位群 U_p 至今仍没有满意的答案，但是库默尔找到 U_p 的一个相当大的子群. 对每个 $i(2\leqslant i\leqslant p-1)$，库默尔发现$\frac{1-\zeta_p^i}{1-\zeta_p}$都是

单位(分圆单位),并且由它们所生成的分圆单位群是 U_p 的指数有限的子群. 利用分圆单位,库默尔对于 $p\nmid xyz$ 的情形(第一种情形)比较容易地推导出矛盾. 但是对 $p|xyz$(第二种)情形,又碰到很大困难,库默尔研究出单位的另一个重要特性(现在叫作库默尔引理),利用这个性质,采用费马的无穷下降法又推出矛盾. 从而最终证明了:若 $\mathbf{Z}[\zeta_p]$是 UFD,则费马猜想对 p 成立. 库默尔在证明他的引理时,创造性地使用了 p-adic 逼近和局部化的思想. 系统的 p-adic 分析和局部化理论是 20 世纪初期由 Hensel 开创的,但是事实上库默尔比 Hensel 早了半个世纪.

事情还未完结. 由于库默尔证明了 $\mathbf{Z}[\zeta_{23}]$不是 UFD,而且他猜想当 $p\geqslant 23$ 时 $\mathbf{Z}[\zeta_p]$均不是 UFD(这个猜想在 1976 年才由 Masley 和 Montgomery 用解析方法证明),所以上述证明只对 $p\leqslant 19$ 适用. 为了进一步证明费马猜想,库默尔于 1847 至 1851 年又产生出新的思想:他发明了一种"理想数",环 $\mathbf{Z}[\zeta_p]$对于数的素分解可能不惟一,但是"理想数"的素分解是惟一的. 一直到 19 世纪后期,狄德金(Dedekind)找到了"理想数"的确切数学含义,它不是一个数,而是环 $\mathbf{Z}[\zeta_p]$的一个特殊子集合,也就是环的理想. 库默尔用"理想数"的惟一分解特性给出费马猜想又一重大突破. 我们用现代语言来叙述这个结果. 环 $\mathbf{Z}[\zeta_p]$的分式理想群对主分式理想群的商群叫作此环的(理想)**类群**. 这是有限阿贝尔群. 此类群的阶 h_p 叫作环 $\mathbf{Z}[\zeta_p]$的(理想)**类数**. $\mathbf{Z}[\zeta_p]$为 UFD 当且仅当 $h_p=1$. 所以前面库默尔证明了:若 $h_p=1$,则费马猜想对 p 成立. 剩下的问题是研究 $h_p\geqslant 2$ 的情形. 库默尔的下一个重要突破是说:

(3) 如果 $p\nmid h_p$,则费马猜想对 p 成立.

例如他算出 $h_{23}=3$,因此费马猜想对 $p=23$ 也成立. 库默尔对于 $p\leqslant 163$ 算出了 h_p 的值(当 $p\leqslant 19$ 时 $h_p=1$)是颇不容易的. 他和欧拉、高斯一样在数论计算方面非常在行,并且他那时已经知道狄里赫利得到的计算分圆域 $\boldsymbol{Q}(\zeta_p)$理想类数 h_p 的解析公式(详见§2.4),但是公式中有一个因子 R_p(叫 regulator)非常难于计算.



图 7 恩斯特·库默尔

所以下一步是希望得到不需要算出 h_p 而能判别 p 是否除尽 h_p 的好办法. 通过对类数公式的仔细分析,他证明了:

(4) 以 h_p^+ 表示 $Q(\zeta_p)$的最大实子域 $Q(\zeta_p+\bar{\zeta}_p)$的理想类数,

则 $h_p^+ | h_p$.

记 $h_p = h_p^+ h_p^-$，则这个结果是说 h_p^- 为整数. 库默尔通过计算发现，h_p^- 随 p 增长得很快，而 h_p^+ 相对来说较小，并且 h_p^- 的计算比较容易. 根据计算 h_p^- 的解析公式，库默尔猜想

$$h_p^- \sim 2p\left(\frac{p}{4\pi^2}\right)^{\frac{p-1}{4}} \quad (\text{当 } p\to\infty \text{时}).$$

这个猜想至今还未能解决. 另一方面，Vandiver 猜想对每个奇素数 p，必然 $p \nmid h_p^+$. 这个猜想至今也未能解决. 但是库默尔证明了：

（5）若 $p | h_p^+$，则 $p | h_p^-$.

由此可知 $p | h_p$ 和 $p | h_p^-$ 是等价的. 这就把问题化成如何判别 p 是否除尽 h_p^-，而 h_p^- 比 h_p 容易计算. 最后，库默尔证明了如下相当满意的结果：

（6）$p | h_p^-$ 当且仅当 p 至少除尽贝努利数 $B_2, B_4, B_6, \cdots, B_{p-3}$ 当中某个的分子. 其中 B_n 由下面展开式的系数所定义（看作关于 t 的形式幂级数）

$$\frac{t}{e^t - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} t^n,$$

数 B_n 是 Bernoulli 在研究级数 $1 + 2^n + 3^n + \cdots + k^n$ 的求和公式时引入的. 由定义可推出计算 B_n 的递归公式：

$$B_0 = 1, \quad B_n = -\frac{1}{n+1}\left(\binom{n+1}{n-1} B_{n-1} + \binom{n+1}{n-2} B_{n-2} + \cdots + \binom{n+1}{1} B_1 + \binom{n+1}{0} B_0\right) \ (n \geqslant 1 \text{ 时}).$$

由此公式可知 B_n 均是有理数. 由定义不难证明当 $n \geqslant 3$ 并且 n 为奇数时 $B_n = 0$，而 $B_1 = -\frac{1}{2}$. 当 n 为偶数时，由递归公式不难算出

$$B_0 = 1,\ B_2 = \frac{1}{6},\ B_4 = -\frac{1}{30},\ B_6 = \frac{1}{42},$$
$$B_8 = -\frac{1}{30},\ B_{10} = \frac{5}{66},\ B_{12} = -\frac{691}{2730}, \cdots$$

所以库默尔结果(4～6)合在一起给出了 $p|h_p$ 的一个初等判别法：$p|h_p \Longleftrightarrow p|h_p^- \Longleftrightarrow p$ 除尽 $B_2, B_4, \cdots, B_{p-3}$ 当中某个的分子. 根据这个判别法，可以算出在 100 以内，除了 37，59 和 67 之外，其他奇素数 p 均满足 $p\nmid h_p$，由结果(3)知费马猜想对这些 p 都是对的. 最后，库默尔又采用进一步的方法证明费马猜想对于 $p=37, 59$ 和 67 也是对的，所以库默尔证明了费马猜想对于不超过 100 的所有奇素数 p 均成立. 这无疑是费马猜想的重大突破. 库默尔把满足 $p\nmid h_p$ 的奇素数叫作正规(regular)的. 那么库默尔结果(3)相当于说：费马猜想对于正规素数 p 成立. 所以从证明费马猜想的角度来说，希望正规素数愈多愈好. 1915 年，Jensen 证明了不正规素数(即 $p|h_p$)有无穷多个. 尽管 Siegel 于 1964 年猜想正规素数与不正规素数的比为 $\sqrt{e}:(\sqrt{e}-1)$，但是至今人们仍不知道正规素数是否有无穷多个.

库默尔数论工作的意义不仅是对费马猜想取得重大突破，他对分圆域所作的深刻研究(特别是对分圆单位群和类数整除性的工作)和高斯对二次域的深刻工作，是代数数论产生的标志. 到目前为止，二次域和分圆域仍旧是研究得最多的两类代数数域.

库默尔的数学研究大致分为三个阶段，在 1842 年之前他研究函数论和微分方程(级数展开和超几何级数的各种性质)，从 1842 至 1859 年他作了 17 年的数论研究. 1859 年之后研究代数几何(库默尔曲面族). 他的数论工作除了上面介绍的之外，还研究过三次高斯和(关于幅角的库默尔猜想至今未被证明)和明显互反律.

参考文献

[1] E. E. Kummer, Collected papers, vol. 1. Springer-Verlag. 1975.

§2.3 狄德金的《代数整数论》(1877年)

继库默尔之后，对于代数数论作了重大贡献的是又一个德国人狄德金(Dedekind,1831—1916).他在19世纪70年代所作的工作是把高斯和库默尔对于二次域和分圆域所作的研究加以一般化和代数化.高斯和库默尔是从研究具体数论问题(不定方程整数解)入手，通过大量实例和计算得到二次域和分圆域的深刻理论结果并用来解决具体数论问题.而狄德金是把他建立的理想论用于数论，用抽象的和代数化的方式讲述数论，重新讲述高斯和库默尔的理论，避免大量的具体计算.狄德金的工作同时是数论和代数发展的一个转折点.他在建立代数化的数论时，建立了近世代数的主要框架：域、环、模和向量空间.这些代数结构和伽罗华的群论一起，形成了抽象代数的核心.他在1877年所写的《代数整数论》一书奠定了经典代数数论的基础.在那个时代，代数数论同时是数论和代数的核心内容.一直到1920年，E. Noether和E. Artin对于环论和模论作了比数论范围更大的推广之后，代数学才从数论中独立出来.即使如此，E. Noether多次说她的工作的许多背景"已经在狄德金之中"了.一直到二战以前学习代数学的年轻人都是从学习代数数论入手的，与现代的情形有很大不同.

近世代数的许多术语都是由狄德金确定的.模(module)来源于高斯研究同余性质时采用的"模m"，域的德文Körper是"身体"的意思，理想来源于库默尔的理想数，至于"环"(ring)据考证是希尔伯特于1897年才采用的，狄德金则使用"整环"(Domain).

狄德金在为狄里赫利《数论讲义》(1871年)一书所写的附录

中首次讲述理想论.他的理论和他在1872年用“狄德金分割”方法定义实数一样,不被当时人所理解.就像E. Artin于1962年在《希尔伯特对现代数学的意义》一文中所说的:“狄德金的工作易懂并且适合于我们今天使用,但在当时是太摩登了.”狄德金的工作特别遭到Kronecker的强烈反对,他反对把无限集合(如理想)作为数学研究对象,反对任何非构造性的抽象研究方式,他甚至不相信无理数.但是Lipschitz于1876年写信支持狄德金把他的理论写成著作,于是有了狄德金于1877年所写的《代数整数论》一书.狄里赫利的《数论讲义》于1879和1894年再版时,狄德金又写了附录,把伽罗华理论结合到他的代数数论中去.1882年,Weber把理想论用于黎曼面的算术理论,把多项式看成是整数,把全纯代数函数看成是代数整数,证明了素理想分解定理,这可以看作是代数几何加以算术化的前驱.希尔伯特也十分欣赏狄德金的工作,他也致力于用代数化方式讲述数论,避免复杂的计算,他用理想论简化了库默尔关于费马猜想的工作.

现在我们介绍狄德金对代数数论的主要贡献.与此同时,我们也介绍经典代数数论的基本概念和结果.

(1) 首先,狄德金确定了(经典)代数数论的基本研究对象和目标.初等数论主要活动场所是有理数域$\boldsymbol{Q}$和它的整数环$\mathbf{Z}$.从17世纪开始,人们研究方程整数(和有理数)解时开始在比$\boldsymbol{Q}$大的范围考虑问题,但是扩大的数都是**代数数**.数α叫作是代数数,是指它是某个次数$\geqslant 1$的有理系数多项式的根.所有代数数形成$\boldsymbol{Q}$的一个扩域$\boldsymbol{Q}^{ac}$,叫作$\boldsymbol{Q}$的**代数闭包**(它是复数域$\boldsymbol{C}$的子域,因为高斯证明了$\boldsymbol{C}$是代数封闭的).但是$\boldsymbol{Q}^{ac}$太大了,它是$\boldsymbol{Q}$的无限次扩张.在研究每个具体不定方程时,通常只需要$\boldsymbol{Q}$的有限次扩张就够了.$\boldsymbol{Q}$的每个扩域K是$\boldsymbol{Q}$上的向量空间,它的维数叫作扩张$K/\boldsymbol{Q}$的**次数**,表示成$[K:\boldsymbol{Q}]$.$\boldsymbol{Q}$的每个有限次扩域K叫作**代数数域**(简称数域),这是代数数论的基本研究对象.二次域和分圆域均是代数数域.代数数论的主要目的是研究代数数域的算术和代数

性质，即算术和代数不变量.

现在可以介绍数域的一批简单的不变量. 数域 K 的次数 $n=[K:\mathbf{Q}]$是 K 的不变量(即它是 K 固有的性质). 可以证明 K 是 $\mathbf{Q}$ 的单扩张，即存在 $\alpha\in K$，使得 $K=\mathbf{Q}(\alpha)$. α 在 $\mathbf{Q}$ 上的最小多项式是指以 α 为根的次数最小的首 1 有理系数的多项式(首 1 是指最高次项系数为 1). 这个多项式 $f(x)$的次数为 $n=[K:\mathbf{Q}]$，并且是由 α 惟一决定的 $\mathbf{Q}$ 上不可约多项式 $f(x)$. 设它的 n 个根为 $\alpha_1,\alpha_2,\cdots,\alpha_n$(其中 α 为它们当中的一个)，这些根彼此不同(因为特征 0 域 $\mathbf{Q}$ 上的不可约多项式没有重根). 由于 $1,\alpha,\alpha^2,\cdots,\alpha^{n-1}$是向量空间 K 的一组 $\mathbf{Q}$-基，我们可以作出 K 到 $\mathbf{C}$ 中的 n 个不同的嵌入(即域的单同态)：

$$\sigma_i:K\to\mathbf{C},\ \sigma_i\left(\sum_{\lambda=0}^{n-1}a_\lambda\alpha^\lambda\right)=\sum_{\lambda=0}^{n-1}a_\lambda\alpha_i^\lambda\qquad(a_\lambda\in\mathbf{Q}),$$

可以证明 K 到 $\mathbf{C}$ 只有这 n 个嵌入 $\sigma_i(1\leqslant i\leqslant n)$. 这也是 K 的不变量(即它不依赖于 α 的选取方式). 设 $\alpha_1,\alpha_2,\cdots,\alpha_{r_1}$是 $f(x)$的实根，其余 $n-r_1$ 个根是 r_2 对共轭复根：$\alpha_{r_1+j}=\bar{\alpha}_{r_1+r_2+j}(1\leqslant j\leqslant r_2)$，$r_1+2r_2=n$. 则 $\sigma_i(K)\subseteq\mathbf{R}(1\leqslant i\leqslant r_1)$，$\sigma_1,\cdots,\sigma_{r_1}$叫作 K 的实嵌入，而其余 r_2 对嵌入叫作 K 的复嵌入：$\sigma_{r_1+j}=\bar{\sigma}_{r_1+r_2+j}(1\leqslant j\leqslant r_2)$. $r_1=r_1(K)$和 $r_2=r_2(K)$也是 K 的不变量(这表明：如果又有 $\beta\in K$ 使得 $K=\mathbf{Q}(\beta)$，则 β 在 $\mathbf{Q}$ 上的最小多项式也是有 r_1 个实根和 r_2 对共轭复根).

(2) 接下来，狄德金把通常整数概念推广到任意代数数域中. 数域 K 中元素 α 叫作代数整数，是指存在首 1 整系数多项式 $f(x)$，使得 $f(\alpha)=0$. 这也相当于 α 在 $\mathbf{Q}$ 上的最小多项式属于 $\mathbf{Z}[x]$. K 中所有代数整数组成的集合记为 O_K. 狄德金的第一个结果为

定理 1 O_K 是整环(叫作 K 的代数整数环).

这也相当于证明：若 α 和 β 为代数整数，则 $\alpha\pm\beta$ 和 $\alpha\beta$ 也如

此. 由定义证明此事并不容易. 1850 年爱森斯坦采用对称函数方法通过复杂的计算证明了此事,而狄德金希望用纯代数方法代替计算,于是他发明了模的概念,证明 α 是代数整数当且仅当 $\mathbf{Z}[\alpha]$ 是有限生成 $\mathbf{Z}$-模,然后用简单而巧妙的推理证明了定理 1.

如果 $K=\mathbf{Q}$,则 $O_K=\mathbf{Z}$,即 $\mathbf{Q}$ 中的代数整数就是通常的整数. 所以狄德金的代数整数概念确实是通常整数的一种推广. 由定义可以决定二次域和分圆域的代数整数环. 对每个整数 $m\geqslant 2$,分圆域 $\mathbf{Q}(\zeta_m)$($\zeta_m=e^{\frac{2\pi i}{m}}$)的代数整数环为 $\mathbf{Z}[\zeta_m]$(就是库默尔的研究对象),对于二次域 $K=\mathbf{Q}(\sqrt{d})$,其中 d 是无平方因子的整数,则

$$O_k=\begin{cases}\mathbf{Z}[\sqrt{d}], & \text{如果 } d\equiv 2,3 \pmod 4,\\ \mathbf{Z}\left[\dfrac{1+\sqrt{d}}{2}\right], & \text{如果 } d\equiv 1 \pmod 4.\end{cases}$$

例如 $\mathbf{Q}(i)$ 的代数整数环为 $\mathbf{Z}[i]$(高斯整数环),而 $Q(\zeta_3)$ 的代数整数环为 $\mathbf{Z}[\zeta_3]$(欧拉证明费马猜想 $n=3$ 情形使用的环).

今后我们经常涉及的是代数整数. 所以我们把代数整数简称为整数,而通常($\mathbf{Z}$ 中)的整数叫作**有理整数**.

数域 K 的(代数)整数环 O_K 是代数数论的重要研究对象. 下面结果是关于 O_K 的加法群结构.

定理 2 O_K 的加法群是秩 $n=[K:\mathbf{Q}]$ 的自由 $\mathbf{Z}$-模(即秩 n 的自由阿贝尔群). 换句话说,存在 n 个整数 $\alpha_1,\cdots,\alpha_n$ 使得 O_k 中每个整数 α 均唯一地表示成

$$\alpha=a_1\alpha_1+\cdots+a_n\alpha_n \quad (a_i\in\mathbf{Z}).$$

这也可以表示成 $O_K=\mathbf{Z}\alpha_1\oplus\mathbf{Z}\alpha_2\oplus\cdots\oplus\mathbf{Z}\alpha_n$(直和).

$\{\alpha_1,\cdots,\alpha_n\}$叫作 O_K(或 K)的一组**整基**,整基并不是 K 的不变量,即可以有多种选取整基的方式. 但是由它可构作出一个不变量. 设 $\sigma_1,\cdots,\sigma_n$ 为数域 K 到 $\mathbf{C}$ 的 n 个嵌入. 对每个 $\alpha\in K$,称 $\sigma_i(\alpha)$

$(1\leqslant i\leqslant n)$为 α 的共轭元素. 而

$$N(\alpha)=\prod_{i=1}^{n}\sigma_i(\alpha),\quad T(\alpha)=\sum_{i=1}^{n}\sigma_i(\alpha)$$

均属于 $\boldsymbol{Q}$,叫作 α 的范和迹. 若$\{\alpha_1,\cdots,\alpha_n\}$是 O_K 的一组整基,则

$$\det(\sigma_i(\alpha_j))^2_{1\leqslant i,j\leqslant n}=\det(T(\alpha_i\alpha_j))_{1\leqslant i,j\leqslant n}$$

是非零有理整数,这个数是 K 的不变量(即与整基的选取方式无关),叫作域 K 的判别式,表示成 $d(K)$. 例如对二次域 $K=\boldsymbol{Q}(\sqrt{d})$,它到 $\boldsymbol{C}$ 有两个嵌入:恒等嵌入 I 和共轭嵌入 σ,其中 $I(\sqrt{d})=\sqrt{d}$, $\sigma(\sqrt{d})=-\sqrt{d}$. 于是对 K 中每个元素 $\alpha=a+b\sqrt{d}\,(a,b\in\boldsymbol{Q})$, $I(\alpha)=\alpha$, $\sigma(\alpha)=a-b\sqrt{d}$. 当 $d\equiv1 \pmod 4$时,$\left\{1,\frac{1+\sqrt{d}}{2}\right\}$是 O_K 的一组整基,因此

$$d(K)=\begin{vmatrix}1 & (1+\sqrt{d})/2\\ 1 & (1-\sqrt{d})/2\end{vmatrix}^2=d.$$

而 $d\equiv2,3\pmod 4$时,$\{1,\sqrt{d}\}$是 O_K 的一组整基,因此

$$d(K)=\begin{vmatrix}1 & \sqrt{d}\\ 1 & -\sqrt{d}\end{vmatrix}^2=4d.$$

(3) 再接下来狄德金开始讲述他的理想论. 整数环 O_K 中的子集合 A 叫作是 O_K 的理想,是指

(a) 若 $\alpha,\beta\in A$,则 $\alpha\pm\beta\in A$(即 A 为加法子群).

(b) 若 $\alpha\in A,\beta\in O_K$,则 $\alpha\beta\in A$(这也可表成 $O_KA=A$).

狄德金定义了理想的运算

$$A+B=\{a+b\,|\,a\in A,b\in B\},$$

$$AB=\left\{\sum_{i=1}^{n}a_ib_i\,|\,a_i\in A,b_i\in B,n\geqslant1\right\}.$$

然后定义理想的整除性:O_K 中理想 A 整除理想 B(表示成 $A|B$),是指存在 O_K 中理想 C,使得 $B=AC$. 易知若 $A|B$,则作为集合必然 $A\supseteq B$. 狄德金在书中详细叙述了他经过长期艰苦的努力,如何

克服困难证明出反过来也成立，即若 A 和 B 是 O_K 的两个理想，$A \supseteq B$，则 $A|B$. 然后他证明了一个重要结果.

定理 3 设 A 是 O_K 的理想，$A \neq (o)$ 或 O_K，则 A 可惟一的表示成有限个素理想的乘积：

$$A = P_1 P_2 \cdots P_k,$$

其中 $P_i(1 \leqslant i \leqslant k)$ 均是 O_K 的素理想（O_K 中理想 P 叫作素理想，是指若 $\alpha, \beta \in O_K, \alpha\beta \in P$，则 $\alpha \in P$ 或 $\beta \in P$）. 惟一性是指若又有 $A = Q_1 Q_2 \cdots Q_r$，其中 $Q_j(1 \leqslant j \leqslant r)$ 均为 O_K 的素理想，则 $k = r$，并且适当调换诸 Q_j 的次序，可使 $P_i = Q_i(1 \leqslant i \leqslant r)$.

这表明，O_K 可能不是元素的惟一因子分解整环（UFD），但 O_K 一定具有理想的素理想惟一因子分解特性，这是库默尔所谓“理想数”惟一因子分解的确切表述. 后人把具有这种性质的任何整环均叫作**狄德金整环**.

在这里自然要问：O_K 中是否有素理想分解的好方法？在有理整数环 $\mathbf{Z}$ 中，每个理想都是主理想 $(n) = n\mathbf{Z}$，并且 (n) 是非零素理想当且仅当 n 是（有理）素数. 现在对素数 p，$\mathbf{Z}$ 中素理想 $p\mathbf{Z}$ 到 O_K 的扩充 pO_K 为 O_K 的理想，但不一定是素理想. 根据定理 3，它应当为 O_K 中素理想的乘积. 我们把相同素理想合并在一起，则

$$pO_K = P_1^{e_1} P_2^{e_2} \cdots P_g^{e_g},$$

其中 $e_i \geqslant 1, P_1, \cdots, P_g$ 是 O_K 的不同素理想. 我们称 g 为 p（在 K 中）的**分解次数**，e_i 叫 p 对 P_i 的**分歧指数**，表示成 $e_i = e(P_i/p)$. 如果 $e_i \geqslant 2$，称 P_i 是**分歧素理想**. 另一方面，如果 $e_1, \cdots, e_g$ 当中至少有一个大于 1，则称 **p 在 K 中分歧**. 最后，在狄德金整环中，每个非零素理想必为极大理想，从而商环 O_K/P_i 是域. 还可证它是特征 p 的有限域，即它是 $\mathbf{Z}/p\mathbf{Z} = \boldsymbol{F}_p$（$p$ 元域）的扩域. 记 $[O_K/P_i : \mathbf{Z}/p\mathbf{Z}] = f_i$，则 O_K/P_i 是 p^{f_i} 元有限域. 我们把 f_i 叫作 p 对 P_i 的**剩余类域次数**，表示成 $f_i = f(P_i/p)$. 上述这些参数有如下的关系：

定理 4 设 K 是 n 次数域(即 $n=[K:\boldsymbol{Q}]$),对于每个素数 p 在 O_K 中的上述素理想分解,则

$$\sum_{i=1}^{g} e(P_i/p)f(P_i/p) = n.$$

下面结果给出 pO_K 素理想分解的一个相当有效的方法.

定理 5 设 α 是数域 K 中的非零整数,则 $\boldsymbol{Z}[\alpha]$是 O_K 的加法子群,并且商群 $O_K/\boldsymbol{Z}[\alpha]$是有限的. 以 $\boldsymbol{F}(x)\in\boldsymbol{Z}[x]$表示 α 在 $\boldsymbol{Q}$ 上的(首 1)最小多项式. 对每个素数 p,如果 $p\nmid|O_K/\boldsymbol{Z}[\alpha]|$(这里 $|S|$表示有限集合 S 的元素个数),$\boldsymbol{F}(x)$在有限域 $\boldsymbol{F}_p$ 上有如下的分解:

$$F(x)\equiv F_1(x)^{e_1}\cdots F_g(x)^{e_g} \pmod p,$$

其中 $F_i(x)(1\leqslant i\leqslant g)$是 $\boldsymbol{F}_p[x]$中的不同首 1 不可约多项式,$e_i\geqslant 1$ $(1\leqslant i\leqslant g)$,则

$$pO_K=P_1^{e_1}P_2^{e_2}\cdots P_g^{e_g},$$

其中 $P_i=(p,F_i(\alpha))$(即是由元素 p 和 $\boldsymbol{F}_i(\alpha)$生成的 O_K 中(素)理想. 而 $f(P_i/p)=\deg F_i(x)$(多项式 $F_i(x)$的次数),$e(P_i/p)=e_i$.

这个定理把素数 p 在 O_K 中的素理想分解和多项式 $F(x)$在有限域 $\boldsymbol{F}_p$ 上的分解联系起来. 这不仅给出素理想分解的一种方法,而且对后人深入思考互反律有重要意义(见本书后面所述).

根据这个定理,容易得到素数在二次域中的分解规律.

例 1 设 $K=\boldsymbol{Q}(\sqrt{d})$,其中 d 是无平方因子的有理整数. p 为有理素数. $d(K)$是 K 的判别式(即当 $d\equiv1 \pmod 4$时 $d(K)=d$,而 $d\equiv2,3 \pmod 4$时 $d(K)=4d$).

(a) 若 p 为奇素数,则

$p\mid d(K)$(即 $p\mid d$)时,$pO_K=P^2$(分歧);

$p \nmid d(K)$时，如果$\left(\frac{d}{p}\right)=1$，则

$pO_K=P_1P_2(P_1\neq P_2)$ （分解）；

如果$\left(\frac{d}{p}\right)=-1$，则 $pO_K=P$ （惰性）.

(b) 若 $p=2$，则

$2|d(K)$时，$2O_K=P^2$ （分歧）；

$2\nmid d(K)$时，必然 $d\equiv 1 \pmod 4$，则

$d\equiv 1 \pmod 8$时，$2O_K=P_1P_2$ （分解）；

$d\equiv 5 \pmod 8$时，$2O_K=P$ （惰性）.

证明 以 $d\equiv 1 \pmod 4$的情形为例. 这时 $d(K)=d$，$O_K=\mathbf{Z}\left[\frac{1+\sqrt{d}}{2}\right]$. 取 $\alpha=\sqrt{d}$，易知$O_K/\sqrt{Z}[\sqrt{d}]$只有两个元素. 从而对每个奇素数 p 可用定理 5. $\alpha=\sqrt{d}$ 的最小多项式为 $F(x)=x^2-d$. 如果 $p|d$，则 $F(x)\equiv x^2 \pmod p$，于是 p 在 K 中分歧：$pO_K=P^2$，$P=(p,\sqrt{d})$. 若$\left(\frac{d}{p}\right)=1$，则存在 $a\in\mathbf{Z}$ 使得 $d^2\equiv a\not\equiv 0 \pmod p$，于是 $F(x)\equiv x^2-a^2\equiv(x-a)(x+a) \pmod p$，于是 p 在 O_K 中分解：$pO_K=P_1P_2$，$P_1=(p,\sqrt{d}+a)$，$P_2=(p,\sqrt{d}-a)$. 如果$\left(\frac{d}{p}\right)=-1$，则 $F(x)$模 p 不可约，于是 $pO_K=P$（即 pO_K 为素理想）.

对于 $p=2$，我们不能取 $\alpha=\sqrt{d}$，而改用 $\alpha=\frac{1+\sqrt{d}}{2}$. 它的最小多项式为 $F(x)=\left(x-\frac{1+\sqrt{d}}{2}\right)\left(x-\frac{1-\sqrt{d}}{2}\right)=x^2-x+\frac{1-d}{4}$ $\left(\text{注意 } d\equiv 1 \pmod 4 \text{ 从而} \frac{1-d}{4}\in\mathbf{Z}\right)$. 若 $d\equiv 1 \pmod 8$，则 $F(x)\equiv x^2-x=x(x-1) \pmod 2$. 于是 2 在 O_K 中分解：$2O_K=P_1P_2$，$P_1=\left(2,\frac{1+\sqrt{d}}{2}\right)$，$P_2=\left(2,\frac{1-\sqrt{d}}{2}\right)$. 若 $d\equiv 5 \pmod 8$，则 $F(x)\equiv x^2+x+1 \pmod 2$不可约，于是 $2O_K=P$. 对 $d\equiv 2,3 \pmod 4$的情形可

类似去做.

特别对高斯数域 $K=\boldsymbol{Q}(i)$，$O_K=\mathbf{Z}[i]$，$d=-1$，$d(k)=-4$. 对奇素数 p，当 $\left(\frac{-1}{p}\right)=1$ 即 $p\equiv1(\bmod 4)$时，$pO_K=P_1P_2$(分解)，当 $\left(\frac{-1}{p}\right)=-1$ 即 $p\equiv3(\bmod 4)$时，$pO_K=P$. 最后 $2O_K=P^2$，$P=(1+i)$. 这就是高斯整数环 $\mathbf{Z}[i]$上元素分解规律的理想论形式.

利用定理 5 也可得到分圆域 $K=\boldsymbol{Q}(\zeta_m)$中素理想分解规律. ζ_m 在 $\boldsymbol{Q}$ 上的最小多项式是 $\varphi(m)$次的"分圆"多项式，

$$\Phi_m(x)=\prod_{\substack{1\leqslant a\leqslant m\\(a,m)=1}}(x-\zeta_m^a)\in\mathbf{Z}[x],$$

即 ζ_m 的全部共轭元素为 $\zeta_m^a(1\leqslant a\leqslant m,(a,m)=1)$(全部 m 次本原单位根). 于是$[K:\boldsymbol{Q}]=\varphi(m)$(欧拉函数).

例 2 设 $K=\boldsymbol{Q}(\zeta_m)$，$p$ 为素数. $m=p^l\cdot n$，其中 $l\geqslant0$ 而$(n,p)=1$. 则

$$pO_K=(P_1\cdots P_g)^{\varphi(p^l)},$$

即所有 $P_i(1\leqslant i\leqslant g)$对 p 的分歧指数均为 $\varphi(p^l)$，剩余类域次数均为 p 模 n 的阶 f(即 f 为满足 $p^f\equiv1(\bmod n)$的最小正整数)，而 $g=\frac{\varphi(n)}{f}$.

从例 1 可知，对于二次域 K，素数 p 分歧当且仅当 $p\mid d(K)$. 狄德金证明了这对任意代数数域都是对的，即

定理 6 素数 p 在数域 K 中分歧，当且仅当 $p\mid d(K)$.

(4) 对每个数域 K，O_K 中非零理想对于乘法形成含幺交换半群(幺元素为 O_K). 而定理 3 是说：这是由(可数无限多个)素理想为基的自由交换半群. 熟知它可扩大成一个交换群，即扩大的元素是 O_K 中非零理想之商 A/B(的等价类). 对于环 O_K，这样的元素有更具体的形式，叫作 K(或 O_K)的分式理想，它是 K 的一个子集

合 M,并且存在 O_K 中非零整数 α,使得 αM 为 O_K 的理想. 以 $I(K)$ 表示 K 的分式理想群,对每个 $o \neq \alpha \in K$,αO_K 为分式理想,叫主分式理想. 全体主分式理想形成 $I(K)$ 的一个子群,表示成 $P(K)$. 两个分式理想 A 和 B 叫作等价的,是指 A/B 为主分式理想. 于是,分式理想等价类数就是商群 $C(K)=I(K)/P(K)$ 的阶. O_K 中的理想今后称作整理想. 这些概念的引入背景是:如何衡量和判别 O_K 是 UFD(即有元素的惟一素因子分解特性)? 我们有

定理 7 对任意代数数域 K,下列三个条件彼此等价:

(1) $I(K)=P(K)$,即每个分式理想都是主分式理想;

(2) O_K 为主理想整环,即每个整理想都是主理想;

(3) O_K 为 UFD(即 O_K 中有元素的惟一素因子分解特性).

继而狄德金证明了又一个重要结果.

定理 8 (类数有限性)对每个代数数域 K,商群 $C(K)=I(K)/P(K)$ 是有限(阿贝尔)群.

$C(K)$ 叫 K 的理想类群,$C(K)$ 中元素叫作一个理想类,每个理想类中两个分式理想相差主理想因子. 群 $C(K)$ 的阶 $h(K)=|C(K)|$ 叫作 K 的理想类数. 定理 7 表明,O_K 为 UFD 当且仅当 $h(K)=1$. 所以群 $C(K)$ 的大小在一定意义下衡量 O_K 与(元素的)唯一素因子分解特性相距多远. 这就是库默尔研究分圆域 $\boldsymbol{Q}(\zeta_p)$ 的类数 h_p 的确切定义. 同时,狄德金把拉格朗日和高斯关于二元二次型的工作用理想论加以改造. 粗糙地说,判别式为 Δ 的二元二次型相当于判别式为 Δ 的二次域 $K=\boldsymbol{Q}(\sqrt{\Delta})$ 的整理想,两个二元二次型在高斯意义下的等价相当于 K 中对应整理想的等价,而高斯定义的二元二次型之间复杂的合成运算恰好对应为整理想的乘积运算(于是合成运算的结合律与交换律在理想论中就变成显

然的事情). 最后，判别式为 Δ 的二元二次型的等价类型 $h(\Delta)$ 就是二次域 K 的理想类数 $h(K)$. 从此以后，研究二元二次型表示整数的问题就可以采用二次域中理想论的语言.

(5) 关于 O_K 的单位群 U_K 的结构，狄里赫利证明了如下的漂亮结果：

定理 9 以 r_1 和 r_2 分别表示数域 K 到 $\boldsymbol{C}$ 的实嵌入个数和复嵌入对的个数，则

$$U_K=W_K\times V_K(\text{直积}),$$

其中 W_K 是 K 中单位根全体(这是有限循环群)，而 V_K 是秩为 $r=r_1+r_2-1$ 的自由交换群. 换句话说，O_K 中存在一组单位 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r$，使得 O_K 中每个单位 μ 均可惟一表示成

$$\mu=\omega\varepsilon_1^{a_1}\varepsilon_2^{a_2}\cdots\varepsilon_r^{a_r},$$

其中 ω 为 K 中的单位根，而 $a_i\in\boldsymbol{Z}(1\leqslant i\leqslant r)$.

$\{\varepsilon_1,\cdots,\varepsilon_r\}$ 叫作 K(或 O_K)的一组**基本单位系**. 基本单位系不是惟一的，但是由它可决定出域 K 的又一个不变量 $R(K)$，叫作 K 的 regulator(定义从略).

例 对于虚二次域 $K=\boldsymbol{Q}(\sqrt{d})(d<0)$，$r_1=0$，$r_2=1$，$r=r_1+r_2-1=0$. 于是 $U_K=W_K$，即 O_K 的单位是 K 中单位根构成的有限群. 当 $d=-1$ 时，$\boldsymbol{Z}[i]$的单位群为$\{\pm1,\pm i\}$，当 $d=-3$ 时，$\boldsymbol{Z}[\zeta_3]=\boldsymbol{Z}\left[\frac{1+\sqrt{-3}}{2}\right]$的单位群为$\{\pm1,\pm\zeta_3,\pm\zeta_3^2\}$. 而对其余虚二次域，$U_K=\{\pm1\}$.

对于实二次域 $K=\boldsymbol{Q}(\sqrt{d})(d>0,\sqrt{d}\notin\boldsymbol{Q})$，$r_1=2$，$r_2=0$，$r=r_1+r_2-1=1$. 于是 K 的基本单位系只含一个单位 ε，而 $W_K=\{\pm1\}$(实单位根只有 ±1). 即实二次域的单位为 $\pm\varepsilon^a(a\in\boldsymbol{Z})$，$\varepsilon$ 叫 K 的**基本单位**.

对于分圆域$K_m=\boldsymbol{Q}(\zeta_m)$，$m\geqslant 3$，$r_1=0$，$r_2=\frac{\varphi(m)}{2}$，$r=r_1+r_2-1=\frac{\varphi(m)}{2}-1$. 当$m$很大时，$r$是很大的正整数，从而$K_m$的单位群是很大的群，决定它的一组基本单位系至今仍是困难的问题(库默尔当年就遇到过这个难题).

以上就是经典代数数论的最基本概念和内容. 核心问题是决定数域理想类群$C(K)$的结构，计算理想类数$h(K)$和决定K的基本单位系. 狄德金对于计算类数$h(K)$作了第一步贡献，他定义了域K的 zeta 函数，在这个基础上后人给出了类数解析公式，于是在代数数论中引进了解析方法.

参 考 文 献

[1] R. Dedekind, Sur la Théorie des Nombres Entiers Algebriques, Gauthier-Villars, 1877. (英译本: Theory of algebraic integers, Cambridge Univ. Press, 1996. 译者 J. Stillwell.)

§2.4 解析方法

解析数论在19世纪中期诞生，是数论史的又一重大事件. 数

论中采用解析方法可以上溯到欧拉. 在 18 世纪中期，欧拉在研究费马数论工作的同时，还热衷于研究各种级数(甚至发散级数)的求和(他对发散级数的研究是想实现他的一种思想：每个发散级数的不同的合理求和方法都应得到同样的结果). 其中一个级数是

$$\zeta(s) = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \cdots + \frac{1}{n^s} + \cdots = \sum_{n=1}^{\infty} n^{-s},$$

这里 s 为实变量. 熟知当 $s>1$ 时此级数收敛，否则级数是发散的. 1736 年，欧拉算出 $\zeta(2)=\frac{\pi^2}{6}$，并且猜想对每个正整数 n，$\zeta(2n)=\pi^{2n}\cdot R$，其中 R 应当是有理数，但是他不能把 R 算出来. 1749 年(1768 年发表)他推出了：当 n 为正整数时

$$\begin{aligned}&\frac{1-2^{n-1}+3^{n-1}-\cdots}{1-2^{-n}+3^{-n}-\cdots}\\ &=\frac{-1\cdot 2\cdot 3\cdot\cdots\cdot(n-1)(2^n-1)}{(2^{n-1}-1)\pi^n}\cos\frac{n\pi}{2}.\end{aligned}$$

并且说这个公式对 n 为任何实数均对(注意上式左边的分子为发散级数). 这个等式左边为 $\zeta(1-n)(1-2^n)/\zeta(n)(1-2^{1-n})$，这实际上相当于一百年后黎曼证明的 $\zeta(s)$ 函数方程! 1737 年，欧拉把级数 $\zeta(s)$ 写成无穷乘积的形式：

$$\begin{aligned}\zeta(s) = \sum_{n=1}^{\infty} n^{-s} &= \prod_p (1 + p^{-s} + p^{-2s} + \cdots + p^{-ms} + \cdots)\\ &= \prod_p (1 - p^{-s})^{-1} \quad (s > 1),\end{aligned}$$

其中无穷乘积过所有素数 p. 不难看出，这个等式相当于正整数 n 的惟一因子分解性质，这就把解析等式和算术性质联系起来. 欧拉用这个等式来证明素数有无穷多个：如果素数只有有限多个，在上式中取 $s=1$，则右边为有限乘积，其值有限，而左边级数发散，这就导出矛盾. 这虽然是一个平凡的例子，但是欧拉这种用解析方法推出数论结果的思想却意义重大. 到了 19 世纪，狄里赫利按照欧拉的思想证明了一个不平凡的数论结果. 设



图 8　Gustav Peter Lejeune Dirichlet 狄里赫利

$$\chi:(\mathbf{Z}/m\mathbf{Z})^{\times}\to \boldsymbol{C}^{\times}$$

是乘法群$(\mathbf{Z}/m\mathbf{Z})^{\times}$到非零复数乘法群的群同态，它也可看成是定义在整数集合 $\mathbf{Z}$ 上周期为 m 的函数（当$(a,m)>1$时，规定 $\chi(a)=0$），后人把 χ 叫作模 m 的**狄里赫利特征**. 他考虑

$$L(s,\chi)=\sum_{n=1}^{\infty}\chi(n)n^{-s}\quad(\mathrm{Re}s>1).$$

后人把它叫作**狄里赫利** L-**函数**. 由于 χ 是群同态，可知 $L(s,\chi)$也有欧拉型的无穷乘积展开

$$L(s,\chi)=\prod_{p}(1-\chi(p)p^{-s})^{-1}.$$

狄里赫利仔细研究了这种复变函数的解析特性，并由此推出：对于任意两个互素的正整数 k 和 l，在算术级数 $(l,l+k,l+2k,\cdots,l+nk,\cdots)$（首项为 l 而公差为 k）中一定存在无限多个素数. 对于一些特殊的 l 和 k，可以有初等的证明方法. 事实上，费马曾经猜想有无穷多个素数 $p\equiv 1 \pmod 4$（这相当于 $l=1,k=4$），而欧拉给出它的初等证明. 但是至今还没有初等方法证明上述狄里赫利算术级数定理的一般性结果. 解析方法在数论中开始显示威力.

解析数论的重大一步是由德国人黎曼（Riemann，1826—1866）做出的. 他只发表了一篇和数论有关的文章，但是却开创了解析数论（韦依（A・Weil）于 1974 年《数论的过去和未来》一文中曾向人们提了一个问题：黎曼是狄里赫利和爱森斯坦的学生，并且与高斯和狄德金关系密切，为什么他没有做数论?）. 黎曼和狄里赫利一样，把 $\zeta(s)$ 和 $L(s,\chi)$ 看成是复变函数. 令 $s=\sigma+it(\sigma,t\in \mathbf{R})$，则级数 $\zeta(s)$（后人称之为黎曼 zeta 函数）在 $\sigma>1$ 时收敛，黎曼的主要结果是（1859 年）.

（1）$\zeta(s)$ 可以解析延拓成整个复平面上的亚纯函数，并且满足下面函数方程：令 $\zeta(s)=\pi^{-s/2}\Gamma(s/2)\zeta(s)$，则 $\zeta(s)=\zeta(1-s)$. 其中 $\Gamma(s)$ 为伽玛函数.

黎曼给出这一结果的两个证明，我们只介绍其中一个的证明思想. 伽玛函数是欧拉定义的：

$$\Gamma(s)=\int_0^{\infty} x^{s-1}e^{-x}\mathrm{d}x \quad (s=\sigma+it,\mathrm{Re}(s)=\sigma),$$

积分在 $\mathrm{Re}(s)>0$ 时收敛. 用分部积分得到 $\Gamma(s+1)=s\Gamma(s)$. 由此可把 $\Gamma(s)$ 解析延拓到整个复平面上，并且 $\Gamma(n+1)=n!$（n 为正整数）. $\Gamma(s)$ 在整个复平面上只有奇点 $s=0,-1,-2,\cdots$，并且均为单极点，留数为 $\operatorname*{res}\limits_{s=-n}\Gamma(s)=(-1)^n/n!$. $\Gamma(s)$ 在复平面中没有零点. 将 $\Gamma(s/2)$ 与 $\zeta(s)$ 合在一起便有

$$\zeta(s)=\pi^{-s/2}\Gamma(s/2)\zeta(s)=\int_0^{\infty}\pi^{-s/2}\sum_{n=1}^{\infty}n^{-s}t^{\frac{s}{2}-1}e^{-t}\mathrm{d}t$$

$$= \int_0^\infty t^{s/2-1}\left(\sum_{n=1}^{\infty} e^{-\pi n^2 t}\right)\mathrm{d}t \quad (\text{作代换 } t \mapsto \pi n^2 t)$$

$$= \int_1^\infty + \int_0^1$$

$$= \int_1^\infty x^{\frac{s}{2}-1}\omega(x)\mathrm{d}x + \int_1^\infty x^{-\frac{s}{2}-1}\omega(x^{-1})\mathrm{d}x, (*)$$

其中 $\omega(x) = \sum_{n=1}^{\infty} e^{-\pi n^2 x}$.

令

$$\theta(x) = \sum_{n=-\infty}^{\infty} e^{-\pi n^2 x} = 1 + 2\omega(x).$$

这是雅柯比使用过的 theta 函数(雅柯比用它来证明:每个正整数都可表成四个整数的平方和). 利用付立业分析中的 Poisson 求和公式,可得到 $\theta(x)$的函数方程 $\theta(x)=\frac{1}{\sqrt{x}}\theta(x^{-1})$,于是 $\omega(x^{-1})=\frac{\sqrt{x}-1}{2}+\sqrt{x}\,\omega(x)$. 将此式代入($*$)式右边,得到

$$\zeta(s) = -\frac{1}{s} - \frac{1}{1-s} + \int_1^\infty x^{\frac{s}{2}-1}\omega(x)\mathrm{d}x + \int_1^\infty x^{\frac{1-s}{2}-1}\omega(x)\mathrm{d}x.$$

注意右边关于 $s \mapsto 1-s$ 是对称的,于是 $\zeta(s)=\zeta(1-s)$. 由于上式右边对于任何复数 s 均有意义,从而 $\zeta(s)$解析延拓到整个复平面上.

由 $\zeta(s)$的函数方程和 $\Gamma(s)$的解析特性不难推出:

(2) $\zeta(s)$在复平面上只有一个奇点 $s=1$,而且是单极点,留数为 1. 另一方面,$\zeta(s)$在 $s=-2,-4,-6,\cdots$等处为单零点. 这也叫 $\zeta(s)$的**平凡零点**. 其余零点均在区域 $0<\mathrm{Re}(s)<1$ 之内,并且这些非平凡零点关于垂直线 $\mathrm{Re}(s)=\frac{1}{2}$是对称的.

接下来,黎曼提出了一个重大的猜想:

(3) $\zeta(s)$的所有非平凡零点都在直线 $\mathrm{Re}(s)=\frac{1}{2}$之上.

关于$\zeta(s)$的这些解析性质都有深刻的数论背景,利用这些解析性质可以研究素数分布和各种数论问题(包括大家熟知的哥德巴赫问题).人们已经从黎曼猜想推出许多好的数论结果,但是黎曼猜想至今未解决.

采用类似的方法,也可把狄里赫利 L-函数 $L(s,\chi)$解析延拓成整个复平面上的解析函数,并且有函数方程把 $L(s,\chi)$和 $L(1-s,\overline{\chi})$联系起来,其中 $\overline{\chi}$ 是 χ 的共轭特性,即 $\overline{\chi}(a)=\overline{\chi(a)}$(复共轭).函数方程中除了出现伽玛函数之外,还出现关于 χ 的高斯和.此外,人们也猜想:$L(s,\chi)$的非平凡零点都在直线 $\mathrm{Re}(s)=\frac{1}{2}$之上,这叫广义黎曼猜想,它也没有解决.

解析方法用于代数数论是由狄德金开始的.对每个数域 K,狄德金引入

$$\zeta_K(s)=\sum_A N(A)^{-s}=\sum_{n=1}^{\infty}a_n n^{-s}\quad(\mathrm{Re}(s)>1),$$

其中 A 过 O_K 的所有非零整理想,而 $N(A)=\left|\frac{O_K}{A}\right|$(叫作理想 A 的**范**),而 a_n 是 O_K 中范为 n 的整理想的个数.这个级数在 $\mathrm{Re}(s)>1$ 时是收敛的,叫作数域 K 的**狄德金 zeta 函数**.由于 $N(A)$是积性函数,即 $N(AB)=N(A)N(B)$,并且每个非零整理想都可唯一表成素理想之乘积,所以 $\zeta_K(s)$也有欧拉乘积展开

$$\zeta_K(s)=\prod_P(1-N(P)^{-s})^{-1}\quad(\mathrm{Re}(s)>1),$$

其中 P 过 O_K 的所有(非零)素理想.

关于 $\zeta_K(s)$的解析延拓和函数方程则是 20 世纪的事情了.20 世纪 20 年代德国人 Hecke 做了这件事,他采用了高维 theta 函数,得到 $\zeta_K(s)=H_K\zeta_K(1-s)$,其中 H_K 是相当复杂的表达式.theta 函数是一类重要的模形式,Hecke 的证明标志模形式理论对

数论的作用，关于模形式我们将在第三章介绍. 由 $\zeta_K(s)$ 的函数方程可以得到：

(1) $\zeta_K(s)$ 只有一个奇点 $s=1$，并且是单极点；

(2) $\zeta_K(s)$ 在 $s=0$ 处为 $r=r_1+r_2-1$ 阶零点，在 $s=-2,-4,-6,\cdots$ 等处为 $r+1$ 阶零点，在 $s=-1,-3,-5,\cdots$ 处为 r_2 阶零点. 这些均叫 $\zeta_K(s)$ 的平凡零点. 其他(非平凡)零点都在区域 $0<\mathrm{Re}(s)<1$ 之中，并且关于直线 $\mathrm{Re}(s)=\frac{1}{2}$ 是对称的.

(3) 人们猜想：对每个数域 K，$\zeta_K(s)$ 的非平凡零点均在直线 $\mathrm{Re}(s)=\frac{1}{2}$ 之上. 这个猜想至今对任何数域都未解决.

问题：$\zeta_K(s)$ 在单极点 $s=1$ 处的留数是什么? 狄里赫利算出了

定理 1 $\zeta_K(s)$ 在 $s=1$ 处的留数为

$$\frac{2^{r_1}(2\pi)^{r_2}R(K)h(K)}{w_K|d(K)|^{1/2}},$$

其中 r_1 和 r_2 为 K 的实嵌入个数和复嵌入对的个数，w_K 为 K 中单位根的个数，而 $d(K)$，$R(K)$ 和 $h(K)$ 分别是 K 的判别式，regulator 和理想类数.

这个结果是令人振奋的. 首先，它把 $\zeta_K(s)$ 的解析特性(在 $s=1$ 处的留数)和数域 K 的众多不变量联系起来. 其次，定理的证明要用到经典代数数论关于整数环、单位群和类群的重要结果. 其中 r_1，r_2，w_K 和 $d(K)$ 通常是容易计算的. $R(K)$ 比较困难，因为它需要知道 K 的基本单位系. 如果我们能用其他办法算出 $\zeta_K(s)$ 在 $s=1$ 处的留数，而且 $R(K)$ 也知道，由上述定理我们便给出计算类数 $h(K)$ 的解析公式. 对于分圆域，狄里赫利解决了第一个困难.

定理 2 对于分圆域 $K=\boldsymbol{Q}(\zeta_m)$，

$$\zeta_K(s) = \prod_{\chi} L(s,\chi^*),$$

其中 χ 过模 m 的所有狄里赫利特征，而 χ^* 是与 χ 相结合的本原特征(关于本原特征的定义从略). 当 χ 为主特征 χ_0 时(主特征指对每个与 m 互素的整数 a，均有 $\chi_0(a)=1$)，$L(s,\chi_0^*)$就是黎曼 zeta 函数 $\zeta(s)$. 而 $\zeta(s)$在 $s=1$ 的留数为 1，于是 $\zeta_K(s)$在 $s=1$ 的留数为

$$\prod_{\chi\neq\chi_0} L(1,\chi^*).$$

换句话说，$\zeta_K(s)$的留数可以用一批狄里赫利 L-函数在 $s=1$ 处的值表达出来. 进而，狄里赫利又给出 $L(1,\chi^*)$一个容易计算的表达式，于是乘积 $R(K)h(K)$便可计算出来(具体公式从略). 所以分圆域类数的计算，需要克服计算 $R(K)$的困难.

库默尔在研究费马猜想时，利用了狄里赫利给出的 $R(K)h(K)$计算公式. 对于 $m=p$ 为素数，以 R_p 和 h_p 表示分圆域 $\boldsymbol{Q}(\zeta_p)$ 的 regulator 和类数，以 R_p^+ 和 h_p^+ 表示极大实子域 $\boldsymbol{Q}(\zeta_p+\overline{\zeta}_p)$的 regulator 和类数，由狄里赫利结果可以计算 R_ph_p 和 $R_p^+h_p^+$. 从而可以计算$\frac{R_ph_p}{R_p^+h_p^+}=h_p^-\ \frac{R_p}{R_p^+}$. 库默尔发现，尽管计算 R_p 和 R_p^+ 非常困难，但是 R_p/R_p^+ 是非常简单的整数 $2^{\frac{p-3}{2}}$. 于是相对类数 h_p^- 可以计算，并且证明了 h_p^- 是整数. 进而，由

$$\begin{aligned}\Gamma(s)\zeta(s) &= \int_0^\infty t^{s-1}e^{-t}\sum_{n=1}^\infty n^{-s}\mathrm{d}t = \int_0^\infty t^{s-1}\sum_{n=1}^\infty e^{-nt}\mathrm{d}t \\ &= \int_0^1 + \int_1^\infty,\end{aligned}$$

易知 $\int_1^\infty$ 是整个复平面上全纯函数，而

$$\begin{aligned}&\int_0^1 t^{s-1}\sum_{n=1}^\infty e^{-nt}\mathrm{d}t \\ &\qquad = \int_0^1 t^{s-1}\frac{dt}{e^t-1} = \int_0^1 t^{s-2}\sum_{n\geqslant 0}\frac{B_n}{n!}t^n\mathrm{d}t\end{aligned}$$

$$= \sum_{n\geqslant 0} \frac{B_n}{n!(n+s-1)}.$$

由于 Bernoulli 数 $B_{2n} \neq 0 (n\geqslant 1)$，可知 $\Gamma(s)\zeta(s)$在 $s=-(2n-1)$ $(n\geqslant 1)$处有单极点，且留数为$\frac{B_{2n}}{(2n)!}$. 另一方面，$\Gamma(s)$在 $s=-(2n-1)$处留数为$-\frac{1}{(2n-1)!}$于是 $\zeta(-2n+1)=-\frac{B_{2n}}{2n}(n\geqslant 1)$. 再由 $\zeta(s)$的函数方程可得到

$$\zeta(2n)=\frac{(-1)^{n+1}(2\pi)^{2n}B_{2n}}{2(2n)!} \quad (n\geqslant 1).$$

这就证明了欧拉的猜想：$\frac{\zeta(2n)}{\pi^{2n}}$是有理数. $\zeta(s)$在特殊整数点的值与 Bernoulli 数发生联系，也是库默尔证明 $p\,|\,h_p$ 可用 Bernoulli 数来判别的一个出发点.

对于分圆域的每个子域 K，采用伽罗华理论和素数在 O_K 中的分解规律，可以相仿地证明 $\zeta_K(s)$是一些狄里赫利 L-函数的乘积，由此也可得到计算 $R(K)h(K)$的解析公式. 二次域是分圆域的子域，而且二次域的 $R(K)$容易计算：当 K 为虚二次域时 $R(K)=1$，当 K 为实二次域时 $R(K)=\log\varepsilon$，其中 $\varepsilon(>1)$为 K 的基本单位. 于是便得到二次域类数的解析公式(具体公式从略).

以上着重介绍了解析方法在计算类数中的应用. 关于解析方法在代数数论中的其他应用我们今后将陆续介绍.

参 考 文 献

[1] R. Dedekind, Mathematische Werke, Vol Ⅰ, Ⅱ, Chelsea, 1968.

[2] L. Dirichlet, R. Dedekind, Vorlesungen uber Zahlentheorie, Chelsea, 1968(原书初版于1893).



图9　希尔伯特

§2.5　希尔伯特的《数论报告》(1897年)

现在我们来到19世纪末期. 1893年,德国数学会委托希尔伯特(Hilbert, 1862—1943)在两年内就代数数论的发展现状写个报

告.两年后,当希尔伯特把“报告”提交给德国数学会时,令所有的人感到惊讶.这哪里是一份报告,而是他本人对代数数论多年来研究成果的总结,也是对一百年来代数数论发展的一个完善的整理,被人们称为“艺术杰作和数学文献中的瑰宝”(相比之下,德国数学会同时委托闵柯夫斯基撰写有理数域上数论的发展现状,两年后他只字未写宣布退出).这本巨著于 1897 年以《代数数域的理论》为书名出版,但数学界一直把它称之为《数论报告》(Zahlbericht).20 世纪的许多著名数学家(E. Artin,H. Hasse,E. Hecke 等)都是通过学习此书开始的数论生涯.

希尔伯特在书的序言中说:数论从它产生的时候起,就以它基础的简洁,概念的清晰和论断的明确有别于其他学科.它有许多结果叙述简单易懂,能够凭直觉相信它,但证明常常很难,但这正是数论的迷人之处,使人“无法从中逃脱”.希尔伯特称赞了前人的成就和对数论的巨大热情,阐述了在他们的努力下数论如何发展到以代数数域为主要研究对象的过程.然后说,他在书中要以统一的观点和合理的逻辑来整理代数数论百年来的理论结果.所有的结果在书中都加以证明,并且在诸多证明中采用最简单或者对数论发展最有启发性的证明.特别致力于用抽象的代数思考方式代替复杂的计算.他声明,书中完全避免历史的论证和关于结果优先权的讨论.

《数论报告》共分五个部分.第一部分是代数数域的一般理论,围绕代数数论的“三个基本柱石:理想的惟一因子分解,单位群结构和类数解析公式”,全面整理了狄德金和狄里赫利等人建立的代数数论基本构架,并且作了深入的研究和发展.最主要的新贡献是:更一般地研究数域扩张 L/K(相对扩张)(狄德金则主要研究 $K=Q$ 的情形),把伽罗华理论和数域扩张更紧密结合起来,建立了精细的分歧理论.第二部分专门介绍数域伽罗华扩张的理论.后三部分则致力于讲述阿贝尔扩张和相对阿贝尔扩张,其中第三部分讲相对二次扩张 $L/K([L:K]=2)$,包括二次域的古典例子,

第四部分讲分圆域，第五部分讲数域的库默尔扩张（即伽罗华扩张 L/K，其伽罗华群为 l 阶循环群，并且 $\zeta_l \in K$）和高次互反律. 他重新整理并极大发展了库默尔结果，他认为库默尔扩张理论及在高次互反律的应用是代数数论的一个高峰，但明确表示他要用理性思维来代替库默尔的复杂计算.

希尔伯特在书中讲述了他本人对代数数论研究的许多新贡献. 除了上述的分歧理论等方面之外，书中花很大篇幅讲述库默尔扩张的种（genus）理论，发明了幂剩余符号和（现在称之为）局部希尔伯特符号来表达"明显"的互反律. 他对于 Kronecker-Weber 定理给了自己的证明（这个定理是说：每个阿贝尔数域都是分圆域的子域）. 基于对相对二次扩张的深刻研究，在书中（和后来在 1899 和 1902 年的继续研究工作中）对于阿贝尔扩张的各种性质进行了猜测，这些猜测在 20 世纪初期得到确认，并推动了类域论的建立. 这里不打算全面介绍希尔伯特这些新结果. 为了今后的需要，我们只简单介绍他关于数域伽罗华扩张的分歧理论.

设 L/K 是数域扩张. L 到 $\boldsymbol{C}$ 中的嵌入 $\sigma: L \to \boldsymbol{C}$ 叫作 K-嵌入，是指 σ 在 K 上的限制是恒等映射，即对每个 $a \in K$，$\sigma(a)=a$. $\sigma(L)$ 叫作 L 的 K-共轭域. 令 $n=[L:K]$，则 L 到 $\boldsymbol{C}$ 共有 n 个不同的 K-嵌入. 如果每个 K-嵌入 σ 都把 L 映到自身，即 $\sigma(L)=L$，则 n 个 K-嵌入均叫作 L 的 K-自同构，并且称 L/K 为**伽罗华扩张**，这 n 个 K-自同构形成群，叫 L/K 的**伽罗华群**，表示成 $\mathrm{Gal}(L/K)$. 例如二次域和分圆域对于 $\boldsymbol{Q}$ 都是伽罗华扩张. 如果 $\mathrm{Gal}(L/K)$ 是阿贝尔群（或循环群），则 L/K 叫作**阿贝尔扩张**（或**循环扩张**）.

设 L/K 是数域扩张，O_K 的每个非零素理想 $\mathscr{P}$ 扩充成 O_L 的理想 $\mathscr{P}O_L$，于是有分解式

$$\mathscr{P}O_L = P_1^{e_1} \cdots P_g^{e_g}, \qquad (*)$$

其中 $P_1, \cdots, P_g$ 为 O_L 的不同素理想，$e_i \geqslant 1 (1 \leqslant i \leqslant g)$. 称 e_i 为 P_i 对 $\mathscr{P}$ 的**分歧指数**，表示成 $e(P_i/\mathscr{P})$. 另一方面，O_L/P_i 和 $O_K/\mathscr{P}$ 均

是有限域，并且前者是后者的扩域，扩张次数$\left[\frac{O_L}{P_i}:\frac{O_K}{\mathscr{P}}\right]$叫作$P_i$对$\mathscr{P}$的**剩余类域次数**，表示成$f(P_i/\mathscr{P})$. 我们也有$\sum_{i=1}^{g}e(P_i/\mathscr{P})f(P_i/\mathscr{P})=[L:K]$.

现在设L/K是伽罗华扩张. 则L的每个K-自同构$\sigma\in \mathrm{Gal}(L/K)$，$\sigma(\mathscr{P})=\mathscr{P}$，于是$\sigma$为$\{P_1,\cdots,P_g\}$的置换，即群$\mathrm{Gal}(L/K)$作用在集合$\{P_1,\cdots,P_g\}$之上. 可以证明这个作用是可迁的，于是将$\mathrm{Gal}(L/K)$作用于分解式(*)，可知$e_1=e_2=\cdots=e_g$，即每个$P_i(1\leqslant i\leqslant g)$对$\mathscr{P}$有相同的分歧指数，类似地推理可知也有相同的剩余类域次数. 于是对伽罗华扩张L/K，便有

$$\mathscr{P}O_K=(P_1\cdots P_g)^e,\quad e=e(P_i/\mathscr{P}),$$

$$f=f(P_i/\mathscr{P})\quad (1\leqslant i\leqslant g),$$

而 $$efg=[L:K](=n=|\mathrm{Gal}(L/K)|).$$

记某个P_i为P，考虑集合

$$G(P)=\{\sigma\in \mathrm{Gal}(L/K):\sigma(P)=P\}.$$

这是P的固定子群. 由$\mathrm{Gal}(L/K)$在$\{P_1,\cdots,P_g\}$上可迁可知$|G(P)|=\frac{n}{g}=ef$. 根据伽罗华扩张的基本定理，$G(P)$对应L/K的一个中间域$L(P)$. $G(P)$和$L(P)$分别叫作P的**分解群**和**分解域**.

记$\bar{L}=O_L/P$，$\bar{K}=\frac{O_K}{\mathscr{P}}$，则$\bar{L}/\bar{K}$是有限域的$f$次扩张. 对每个$\alpha\in O_L$，以$\bar{\alpha}$表示$\alpha$在$\bar{L}$中的象. 以$\bar{G}$表示伽罗华群$\mathrm{Gal}(\bar{L}/\bar{K})$(这是由$\tau$生成的$f$阶循环群，其中$\tau(a)=a^{|\bar{K}|}$(对每个$a\in\bar{L}$)). 现在对每个$\sigma\in G(P)$，由$\sigma(P)=P$知道可以定义一个映射：

$$\bar{\sigma}:\bar{L}=\frac{O_L}{P}\to\bar{L}=\frac{O_L}{P},\quad \bar{\sigma}(\bar{\alpha})=\overline{\sigma(\alpha)}(\text{对每个 }\alpha\in O_L).$$

可直接验证这是域$\bar{L}$的自同构，并且把$\bar{K}$中每个元素均保持不动，于是$\bar{\sigma}\in\bar{G}$. 由此我们得到一个新的映射

$$\varphi: G(P) \to \bar{G}, \quad \varphi(\sigma) = \bar{\sigma}.$$

可以证明这是群的满同态，它的核记为 $G(P)'$.

$$\sigma \in \ker(\varphi) \Leftrightarrow \bar{\sigma} = Id' \Leftrightarrow \bar{\sigma}(\bar{\alpha}) = \bar{\alpha} \quad (\text{对每个 } \alpha \in O_L)$$

$$\Leftrightarrow \overline{\sigma(\alpha)} = \bar{\alpha}(\text{对每个 } \alpha \in O_L)$$

$$\Leftrightarrow \sigma(\alpha) \equiv \alpha (\text{mod } P)(\text{对每个 } \alpha \in O_L).$$

即 $G(P)' = \ker(\varphi) = \{\sigma \in G(P): \sigma(\alpha) \equiv \alpha (\text{mod } P)$，对每个 $\alpha \in O_L\}$. 而 $G(P)/G(P)'$ 同构于 $\bar{G}$. 于是 $|G(P)'| = \frac{|G(P)|}{|\bar{G}|} = \frac{ef}{f} = e$. 以 $L(P)'$ 表示与子群 $G(P)'$ 对应的中间域，$G(P)'$ 和 $L(P)'$ 分别叫作 P 的**惰性群**和**惰性域**. 于是对于 O_L 的素理想 P，我们有如下的群和域之间伽罗华对应的图表.

素理想分解	域		群	剩余类域	
P	L		$\{1\}$	$\bar{L}$	
$\vert$	$\vert$	e	$\vert$	$\Vert$	
P'	$L(P)'$		$G(P)'$	$\overline{L(P)'}$	
$\vert$	$\vert$	f	$\vert$	$\vert$	f
P_D	$L(P)$		$G(P)$	$\overline{L(P)}$	
$\vert$	$\vert$	g	$\vert$	$\Vert$	
$\mathscr{P}$	K		$\text{Gal}(L/K)$	$\bar{K}$	

设 O'，O_D 分别为数域 $L(P)'$ 和 $L(P)$ 的整数环，则 $P \cap O' = P'$ 和 $P' \cap O_D = P_D$ 分别为 O' 和 O_D 的素理想，而 $P_D \cap O_K = \mathscr{P}$. 记 $\overline{L(P)'} = \frac{O'}{P'}$，$\overline{L(P)} = \frac{O_D}{P_D}$，则我们有图表最右方的有限域扩张. 下面的定理把 $\mathscr{P}$ 在 L 中的分解过程作了细分.

定理 (1) $e(P_D/\mathscr{P}) = f(P_D/\mathscr{P}) = 1$，并且当 $L(P)/K$ 是伽罗华扩张(即 $G(P)$ 是 $\text{Gal}(L/K)$ 的正规子群)时，$\mathscr{P}$ 在 $L(P)$ 中分解成 $g(=[L(P):K])$ 个不同素理想的乘积(完全分裂).

(2) $f(P'/P_D) = f$，$e(P'/P_D) = 1$，即 $P_D O' = P'$(惰性).

(3) $f(P/P')=1, e(P/P')=e$，即 $P'O_L=P^e$（完全分歧）.

(4) 对每个 $\sigma\in \mathrm{Gal}(L/K)$，则有

$L(\sigma(P))=\sigma L(P), L(\sigma(P))'=\sigma L(P)'$,

$G(\sigma(P))=\sigma G(P)\sigma^{-1}, G(\sigma(P))'=\sigma G(P)'\sigma^{-1}$.

特别若 L/K 是阿贝尔扩张，即 $\mathrm{Gal}(L/K)$ 是阿贝尔群，则 $G(\sigma(P))=G(P), G(\sigma(P))'=G(P)'$. 由伽罗华对应可知 $L(\sigma(P))=L(P), L(\sigma(P))'=L(P)'$. 换句话说，$\mathscr{P}$ 在 O_L 中的 g 个素理想因子 $P_i(1\leqslant i\leqslant g)$ 具有同样的分解群、分解域、惰性群和惰性域，它们可分别记为 $G(\mathscr{P}), L(\mathscr{P}), G(\mathscr{P})'$ 和 $L(\mathscr{P})'$.

当 L/K 是阿贝尔扩张时，$G(P)$ 必为 $G(L/K)$ 的正规子群. 所以 $\mathscr{P}$ 的分解可以有更清晰的图画：$\mathscr{P}$ 在分解域 $L(\mathscr{P})$ 中完全分解成 g 个不同素理想之积 $P_{D,1}\cdots P_{D,g}$. 每个 $P_{D,i}$ 在惰性域 $L(\mathscr{P})'$ 的整数环中的扩充仍为素理想. $\mathscr{P}'_i(1\leqslant i\leqslant g)$. 而每个 $\mathscr{P}'_i$ 在 L 中完全分歧：$\mathscr{P}'_iO_L=P_i^e$. 于是最后得到

$$\mathscr{P}=P_{D,1}\cdots P_{D,g}=\mathscr{P}'_1\cdots\mathscr{P}'_g=P_1^e\cdots P_g^e.$$

(5) 若 $e=1$，即 $\mathscr{P}$ 在 L 中不分歧，则 $G(P)'=\{1\}$. 于是分解群 $G(P)$ 和 $\overline{G}$ 同构，即 $G(P)$ 是 f 阶循环群. 由于 $\overline{G}$ 是由 τ 生成的，τ 在 $G(P)$ 中的对应元素记为 $\left(\frac{L/K}{P}\right)$，叫作关于 P 的 Frobenius **自同构**. 则 $G(P)$ 就是由 $\left(\frac{L/K}{P}\right)$ 生成的 f 阶循环群. 由于 τ 把 $\overline{L}$ 中元素 $\bar{\alpha}(\alpha\in O_L)$ 映成 $\bar{\alpha}^{|K|}$，可知关于 P 的 Frobenius 自同构可以由下式来刻划：

$$\left(\frac{L/K}{P}\right)\alpha\equiv\alpha^{|K|}(\mathrm{mod}\ P)（对每个 \alpha\in O_L）.$$

关于 Frobenius 自同构有如下性质：

(a) 对每个 $\sigma\in\mathrm{Gal}(L/K)$，$\left(\frac{L/K}{\sigma(P)}\right)=\sigma\left(\frac{L/K}{P}\right)\sigma^{-1}$. 特别当 L/K 是阿贝尔扩张时，$\left(\frac{L/K}{\sigma(P)}\right)=\left(\frac{L/K}{P}\right)$. 从而关于 $\mathscr{P}$ 的每个素

理想因子 $P_i(1\leqslant i\leqslant g)$有同样的 Frobenius 自同构，记作$\left(\frac{L/K}{\mathscr{P}}\right)$.

(b) 若 E 为 L/K 的中间域，则 L/E 也是伽罗华扩张. 记 $P_E=P\cap O_E$，则$\left(\frac{L/E}{P}\right)=\left(\frac{L/K}{P}\right)^{f(P_E/\mathscr{P})}$.

(c) 如果 E/K 也是伽罗华扩张，则$\left(\frac{E/K}{P_E}\right)=\left(\frac{L/K}{P}\right)$在 E 上的限制.

现在我们举分圆域和它的子域的例子，以解释上述概念和结果.

例 1 取 $L=\boldsymbol{Q}(\zeta_7)$，$K=\boldsymbol{Q}$. 则 L/K 是伽罗华扩张. 以下简记 $\zeta=\zeta_7$，则 ζ 在 $\boldsymbol{Q}$ 上的最小多项式为 $f(x)=\frac{x^7-1}{x-1}=x^6+x^5+x^4+x^3+x^2+x+1$，于是$[L:\boldsymbol{Q}]=6$. ζ 的 6 个共轭元素为 $\zeta^i(1\leqslant i\leqslant 6)$. L 的每个自同构有形式 $\sigma_i(\zeta)=\zeta^i$. 而 $\mathrm{Gal}(L/Q)=\{\sigma_i\,|\,1\leqslant i\leqslant 6\}$. 由于

$$\sigma_i\sigma_j(\zeta)=\sigma_i(\zeta^j)=\sigma_i(\zeta)^j=\zeta^{ij}=\sigma_{ij}(\zeta),$$

即 $\sigma_i\sigma_j=\sigma_{ij}$，于是 $\mathrm{Gal}(L/\boldsymbol{Q})$同构于$(\boldsymbol{Z}/7\boldsymbol{Z})^\times$($\sigma_i$ 对应于 $i\pmod 7$)，从而 L/Q 是阿贝尔扩张. 现在研究每个素数 p 在 $O_L=\boldsymbol{Z}[\zeta]$中的素理想分解规律.

$$\begin{array}{ccccc} L=\boldsymbol{Q}(\zeta_7) & & \{\sigma_1\} & \cong & \{1\} \\ | & & | & & | \\ \boldsymbol{Q} & & \mathrm{Gal}(L/\boldsymbol{Q}) & \cong & (\boldsymbol{Z}/7\boldsymbol{Z})^\times \end{array}$$

可以算出域 L 的判别式为 $d(L)=-7^6$. 根据狄德金的判别式定理(§2.3 定理 6)，当 $p\neq 7$ 时，p 在 L 中不分歧. 于是 $pO_L=P_1\cdots P_g$，$f=f(P_i/p)\quad(1\leqslant i\leqslant g)$，$fg=[L:\boldsymbol{Q}]=6$. 而 f 是 Frobenius 自同构$\left(\frac{L/\boldsymbol{Q}}{p}\right)$的阶. 它由下面公式来刻划：

$$\left(\frac{L/\boldsymbol{Q}}{p}\right)\alpha\equiv\alpha^p\pmod{P_1}\quad(\text{对每个 }\alpha\in O_L=\boldsymbol{Z}[\zeta]).$$

特别地$\left(\frac{L/\boldsymbol{Q}}{p}\right)\zeta\equiv\zeta^p(\mathrm{mod}\ P_1)$. 令$\left(\frac{L/\boldsymbol{Q}}{p}\right)=\sigma_t$,则$\sigma_t(\zeta)=\zeta^t$,于是$\zeta^t\equiv\zeta^p(\mathrm{mod}\ P_1)$. 由此可推出$t\equiv p(\mathrm{mod}\ 7)$,于是我们决定出$\left(\frac{L/\boldsymbol{Q}}{p}\right)=\sigma_p$. 而$f$为$\sigma_p$的阶,又相当于$p$模7的阶. 由于$(\mathbf{Z}/7\mathbf{Z})^\times$中1阶元素只有1,2阶元素有6,3阶元素有2和4,6阶元素有3和5. 于是

$$pO_L=\begin{cases}P_1\cdots P_6, \text{若 } p\equiv1(\mathrm{mod}\ 7)(\text{完全分解});\\ P_1P_2P_3, \text{若 } p\equiv6(\mathrm{mod}\ 7);\\ P_1P_2, \quad \text{若 } p\equiv2,4(\mathrm{mod}\ 7);\\ P, \qquad \text{若 } p\equiv3,5(\mathrm{mod}\ 7)(\text{惰性}).\end{cases}$$

当$p=7$时,可证明它完全分歧:$7O_L=P^6$,$P=(1-\zeta)O_L$.

由于$O_L=\mathbf{Z}[\zeta]$,利用§2.3定理5,可以决定$\mathbf{Z}[x]$中不可约多项式$f(x)=x^6+x^5+x^4+x^3+x^2+x+1$在$\boldsymbol{F}_p[x]$中如何分解. 即

当$p\equiv1(\mathrm{mod}\ 7)$时$f(x)$在$\boldsymbol{F}_p[x]$中分解成6个一次多项式之积. 当$p\equiv6(\mathrm{mod}\ 7)$时$f(x)$在$\boldsymbol{F}_p[x]$中分解成3个二次不可约多项式之积. 当$p\equiv2,4(\mathrm{mod}\ 7)$时,$f(x)$在$\boldsymbol{F}_p[x]$中分解成2个三次不可约多项式之积. 而当$p\equiv3,5(\mathrm{mod}\ 7)$时,$f(x)$在$\boldsymbol{F}_p[x]$中仍不可约. 最后,$f(x)$在$\boldsymbol{F}_7[x]$中是一次多项式$(x-1)$的6次方.

例2 考虑L的最大实子域$K=\boldsymbol{Q}(\zeta+\bar{\zeta})$. 则$K/\boldsymbol{Q}$也是阿贝尔扩张. K的固定子群为$\{\sigma_1,\sigma_{-1}\}$,

$$\begin{array}{cccc}L=\boldsymbol{Q}(\zeta_7) & \{\sigma_1\} & \cong & \{1\}\\ | & | & & |\\ K=\boldsymbol{Q}(\zeta_7+\bar{\zeta}_7) & \{\sigma_1,\sigma_{-1}\} & \cong & (\mathbf{Z}/7\mathbf{Z})^\times/(\pm1)\\ | & | & & |\\ \boldsymbol{Q} & \mathrm{Gal}(L/Q) & \cong & (\mathbf{Z}/7\mathbf{Z})^\times\end{array}$$

于是$\mathrm{Gal}(K/Q)=\frac{\mathrm{Gal}(L/\boldsymbol{Q})}{\{\sigma_1,\sigma_{-1}\}}\cong\frac{(\mathbf{Z}/7\mathbf{Z})^\times}{(\pm1)}$. 当$p=7$时,由于7在$L$中完全分歧,可知在$K$中也完全分歧,即$7O_K=P^3(3=[K:\boldsymbol{Q}])$.

当 $p \neq 7$ 时，p 在 L 中不分歧，可知在 K 中也不分歧，即

$$pO_K = P_1 \cdots P_g,\ f = f(P_i/p)(1 \leqslant i \leqslant g),\ gf = 3.$$

f 为 $\left(\dfrac{K/Q}{p}\right)$ 的阶. 由于 $\left(\dfrac{K/Q}{p}\right)$ 是 $\left(\dfrac{L/Q}{p}\right) = \sigma_p$ 在 K 中的限制，从而 f 相当于 σ_p 在 $\mathrm{Gal}(K/\boldsymbol{Q})$ 中的阶，即相当于 p 在 $(\mathbf{Z}/7\mathbf{Z})^{\times}/(\pm 1)$ 中的阶，也就是说，f 为最小正整数使得 $p^f \equiv \pm 1 \pmod 7$. 易知当 $p \equiv \pm 1 \pmod 7$ 时 $f = 1$，否则 $f = 3$. 于是

$$pO_K = P_1P_2P_3,\ 若\ p \equiv \pm 1 \pmod 7.$$

$$pO_K = P,\ 若\ p \equiv \pm 2,\ \pm 3 \pmod 7.$$

$K = \boldsymbol{Q}(\zeta + \bar{\zeta})$ 中元素 $\alpha = \zeta + \bar{\zeta}$ 的共轭元素为 $\alpha_1 = \alpha, \alpha_2 = \zeta^2 + \bar{\zeta}^2$ 和 $\alpha_3 = \zeta^3 + \bar{\zeta}^3$. 从而 α 在 $\boldsymbol{Q}$ 上的最小多项式为 $f(x) = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3) = x^3 + x^2 - 2x - 1$. 可以验证 $O_K = \mathbf{Z}[\alpha]$，于是：多项式 $f(x) = x^3 + x^2 - 2x - 1$ 在 $\boldsymbol{F}_7[x]$ 中为一次多项式的三次方. 当 $p \equiv \pm 1 \pmod 7$ 时，$f(x)$ 在 $\boldsymbol{F}_p[x]$ 中为三个不同的一次多项式相乘，而 $p \equiv \pm 2, \pm 3 \pmod 7$ 时，$f(x)$ 在 $\boldsymbol{F}_p[x]$ 中不可约.

以上两个例子表明，素数 p 在 $\boldsymbol{Q}(\zeta_7)$ 或它的子域 $\boldsymbol{Q}(\zeta_7 + \bar{\zeta}_7)$ 中的分解模式（即 e, f, g 的值）只依赖于 p 所在的模 7 同余类. 这件事可以看成是分圆域和它子域中的互反律，我们以后再解释它的理由.

以上我们对希尔伯特的《数论报告》作了简要的介绍. 对于这本代数数论的精品，数论学家们仍然有不同的评价和意见. 最厉害的批评性意见来自韦依（A. Weil，1906—1998）. 他在为库默尔全集所写的前言中说："在库默尔去世之后，希尔伯特统治德国数学界多年. 他的著名的数论报告有一半以上的内容（例如第四和第五部分）基本上是库默尔的工作，只有很小的非本质性改进. 缺乏对前人数学风格的尊重，没有介绍库默尔采用的 p-adic 分析，对库默尔工作也引用得不充分. "《数论报告》出版后的第一代读者大多对此书表示赞赏. E. Artin 在 1962 年纪念希尔伯特百年诞辰的演

讲中称赞希尔伯特对库默尔工作的简化. 而 Hasse 也是如此,认为希尔伯特用更好的证明代替了库默尔许多复杂而不透彻的推导. Hasse 认为数论发展中有两种风格不同的传统,高斯和库默尔的传统以构造性和表达明显为目标,后来由 Kronecker 和 Hensel 继承. 另一个是狄德金和希尔伯特传统,注重抽象和概念化的理解. 他赞成两种风格要"有机地平衡".

从 19 世纪开始,世界数论中心从法国转到代数数论和解析数论诞生地,一直维持到二战时期.

参 考 文 献

[1] D. Hilbert, Zahlbericht, Die Theorie der algebraischen Zahlkörper, 1897. 英译本 The Theory of algebraic number fields, Springer-Verlag, 1998.

第三章　近代代数数论
(1900—1967)

“谁不想揭开未来的面纱，瞻望一下我们科学的前景和下世纪科学发展的秘密？什么样的特别的目标将引导未来的一代并激发出主流数学的新精神？在广阔而富饶的数学田野上，下一世纪将揭示出什么样的新方法和新成果？”

这是希尔伯特于 1900 年在第二届世界数学家大会(巴黎)上所作《数学问题》报告的开场白. 他接着讲述对数学发展的理解和观点，其中有一段专门论述代数数论.

“众所周知，费马猜测不定方程 $x^n+y^n=z^n(n\geqslant 3)$ 没有非平凡的整数解. 人们对证明这个猜想所作的努力，是思维功能的一个精彩的例子. 这个很特别并且看来似乎不重要的问题，可以发展成科学. 库默尔在研究费马问题中引入理想数，并且对于分圆域发现了理想数分解成理想素数的惟一性. 这个规律今天已被狄德金和 Kronecker 推广到任意代数数域之中，成为现代数论的中心位置，其意义已超出了数论的界限，走到代数学和函数论的王国中.”

站在新旧世纪交替的门槛上，希尔伯特向人们提出了著名的 23 个数学问题. 在其中的 6 个数论问题中，有 4 个属于代数数论. 这些问题对于推动 20 世纪的代数数论发展起了相当大的作用. 本章我们以这 4 个希尔伯特问题为线索，介绍 20 世纪前 70 年代数数论研究中创造的新思想和新方法，得到的新成果，以及由此产生的代数数论新的重要分支. 我们也将按历史的顺序展示代数数论逐渐与数学其他学科相互交织、渗透和促进的生动场面.

§ 3.1　类域论(20 世纪 20 年代)

希尔伯特第 12 个问题是:Kronecker-Weber 定理如何推广到任意代数数域上?

Kronecker-Weber 定理是说:每个阿贝尔数域都是(某个)分圆域的子域.我们知道,分圆域 $\boldsymbol{Q}(\zeta_m)(m\geqslant 3)$是阿贝尔域,即它是 $\boldsymbol{Q}$ 的伽罗华扩张,并且伽罗华群是同构于$(\mathbf{Z}/m\mathbf{Z})^{\times}$ 的(有限)阿贝尔群.于是分圆域的每个子域都是阿贝尔域,因为其伽罗华群为$(\mathbf{Z}/m\mathbf{Z})^{\times}$ 的某个商群.19 世纪末期,Kronecker 和 Weber 证明了反过来也成立(确切地说,由前者于 1877 年猜测而由后者于 1887 年完成证明).所以 $\boldsymbol{Q}$ 的最大阿贝尔扩张(表示成 $\boldsymbol{Q}^{ab}$)就是所有分圆域的合成:

$$\begin{aligned}\boldsymbol{Q}^{ab} &= \bigcup_{m\geqslant 3}\boldsymbol{Q}(\zeta_m) = \boldsymbol{Q}(\zeta_3,\zeta_4,\zeta_5,\cdots)\\ &= \boldsymbol{Q}(e^{\frac{2\pi i}{m}}:m=3,4,5,\cdots)\\ &= \boldsymbol{Q}(e^{2\pi i\alpha}:\alpha\in\boldsymbol{Q}).\end{aligned}$$

$\boldsymbol{Q}^{ab}$是 $\boldsymbol{Q}$ 的无限次(阿贝尔)扩张,是把周期为 1 的指数函数 $f(x)=e^{2\pi ix}(x\in\boldsymbol{R})$在有理数上的全部取值添加在 $\boldsymbol{Q}$ 上而得到的扩域.

对于任意代数数域 K,如果 L_1 和 L_2 是 K 的两个有限阿贝尔扩域,则易证它们的合成 L_1L_2 也是 K 的有限阿贝尔扩域.所以 K 的最大阿贝尔扩张 K^{ab}一定存在,它就是 K 的所有(无限多个)有限阿贝尔扩域的合成.希尔伯特第 12 个问题是问:如何像 $\boldsymbol{Q}^{ab}$那样用明显的方式把 K^{ab}构作出来,是否有某些特殊的函数,使得把这些函数的一些特殊取值添加在 K 上就能得到 K^{ab}?

这是一个充满着秘密的探索性问题.它本质上是要对 K 的所有有限阿贝尔扩张作深入的了解.这些扩张显然是由 K 所决定的,所以要弄清如何由 K 所固有的性质(哪些性质?)把 K 的所有有限阿贝尔扩张刻划出来.

作为第一步,首先要明显构作出数域 K 充分多的有限阿贝尔扩张.19 世纪库默尔对 Q 构作出相当多的阿贝尔扩张(分圆域).19 世纪末期 Kronecker 和 Weber 证明这足可以得到 Q 的最大 Abel扩张.1883 年～1890 年,Kronecker 研究模函数和(双周期)椭圆函数,发现这些函数在"分点"上的值添加到虚二次域 K 上,可以得到 K 的许多有限阿贝尔扩张.1880 年他猜想将这些有限阿贝尔扩张合并起来就是 K 的最大阿贝尔扩张,他把这个猜想称作"青春之梦"(Jugendtraum).这个猜想曾在 1908 和 1914 年由 Weber 和 Fueter 部分地被解决.一直到 1920 年在德国哥丁根大学工作的日本数学家高木贞治(Takagi)建立了类域论,这个"青春之梦"才完全得以实现.对于 Q 和虚二次域以外的代数数域 K,具体构作 K^{ab}的问题至今未能解决.

类域论是研究数域有限阿贝尔扩张的理论,它完全是在希尔伯特第 12 问题的推动下建立起来的.希尔伯特在《数论报告》和后来的研究中,得到数域阿贝尔扩张的许多性质并作了不少猜测,而高木准确地找到了数域 K 的一批对象,用它们一一对应着 K 的全部有限阿贝尔扩张.这批对象是 K 的理想类群的推广,叫作**广义类群**.

现在我们介绍类域论最基本的内容.类域论从它产生以来,叙述方式和证明方法不断地更新.高木于 1920 年创建类域论时,证明综合采用了代数和解析方法,1927 年 Artin 引入了现在称之为 Artin 映射,把类域论的叙述方式进行了改造(Artin 于 40 年代到美国之后,多年来在大学里讲授类域论,不断改变类域论的面貌).日本数学家和 Artin-Tate 随后又采取了以上同调理论为主的叙述和证明方式.在 Chevalley 建立 Idèle 语言之后,类域论又用这

种语言重新加以叙述. 最后，韦依在《数论基础》一书中，对类域论又使用单代数的方式. 我们现在介绍的是采用 Artin 映射的一种容易理解的方式.

设 K 为数域. $M_0=P_1^{e_1}\cdots P_g^{e_g}$ 为 O_K 的非零整理想，其中 $e_i\geqslant 1$，$P_i(1\leqslant i\leqslant g)$ 为 O_K 的不同素理想. 我们也把 O_K 的每个非零素理想叫作 K 的一个有限位. 对应 K 到 $\boldsymbol{C}$ 的每个实嵌入 σ，我们相应有 K 的一个实位（无限位）∞. 设 M_∞ 是 K 的一些不同实位的（形式）乘积：$M_\infty=\infty_1\cdots\infty_l(0\leqslant l\leqslant r_1)$. 我们把 $M=M_0M_\infty$ 叫作 K 的一个除子，有限位 P_i 和实位 ∞_j 都叫作 M 的因子. 我们在下节讲赋值理论时，再解释位和除子这些概念.

K 的一个分式理想 A 叫作与 M_0 互素，是指 A 可以写成两个整理想 B 和 C 之商，使得 B 和 C 均与 M_0 互素. 以 $I_M(K)$ 表示 K 中与 M_0 互素的所有分式理想（对乘法）构成的群，以 $P_M(K)$ 表示由主分式理想 αO_K 形成的群，其中 $\alpha\in K^\times$ 满足以下两个条件：

(1) $\alpha\equiv 1\pmod{M_0}$，意思是：对每个 $i(1\leqslant i\leqslant g)$，$P_i^{e_i}\mid(\alpha-1)O_K$.

(2) 如果实位 ∞ 为 M_∞ 的因子，而 ∞ 对应于 K 的实嵌入 σ，则 $\sigma(\alpha)>0$.

不难证明 $P_M(K)$ 是 $I_M(K)$ 的子群，也可以证明商群 $C_M(K)$ 是有限（阿贝尔）群，叫作是 K 的模 M 广义类群. 当 $M\doteq 1$（即 $M_0=(1)=O_K$，$M_\infty=1$）时，M 没有任何位作为因子，$P_1(K)$ 和 $I_1(K)$ 就是 K 的分式理想群 $P(K)$ 和主分式理想群 $I(K)$，从而 $C_1(K)$ 就是通常的理想类群 $C(K)=P(K)/I(K)$. 所以 $C_M(K)$ 是理想类群的一种推广.

现在我们用 $K=\boldsymbol{Q}$ 来解释这些概念. $O_K=\boldsymbol{Z}$ 的非零理想都是主理想 $(m)=m\boldsymbol{Z}$，其中 m 为正整数，所以 $M_0=(m)$，我们也简写成 $M_0=m$. 另一方面，$\boldsymbol{Q}$ 到 $\boldsymbol{C}$ 只有一个嵌入（恒等嵌入），从而 $\boldsymbol{Q}$ 只有一个实位 ∞. 所以 M_∞ 只有 ∞ 和 1 两种可能. 于是 M 有形式 m

和 $m\cdot\infty$. 而 m 的素因子 p 相当于 M_0 的素理想因子(或有限位因子).

设 $M=m$. 由于 Q 的分式理想均为主分式理想，可知

$$I_m(Q)=\left\{\left(\frac{a}{b}\right):a\text{ 和 }b\text{ 为正整数},(a,b)=(a,m)=(b,m)=1\right\},$$

$$P_m(Q)=\left\{(\alpha)=(-\alpha):\ \alpha=\frac{a}{b},\ a\text{ 和 }b\text{ 为正整数},\ \frac{a}{b}\equiv 1\ (\mathrm{mod}\ m)\right\}$$

$$=\left\{\left(\frac{a}{b}\right):a\text{ 和 }b\text{ 为正整数},\ (a,b)=1,\frac{a}{b}\equiv\pm 1(\mathrm{mod}\ m)\right\}.$$

不难证明：$C_m(Q)=I_m(Q)/P_m(Q)\cong(\mathbf{Z}/m\mathbf{Z})^\times/(\pm 1)$. 我们知道右边正则同构于 $Q(\zeta_m+\bar{\zeta}_m)$ 对 Q 的伽罗华群，同构为

$$(\mathbf{Z}/m\mathbf{Z})^\times/(\pm 1)\xrightarrow{\sim}\mathrm{Gal}(Q(\zeta_m+\bar{\zeta}_m)/Q),$$

$$P\mapsto\sigma_P=\left(\frac{Q(\zeta_m+\bar{\zeta}_m)/Q}{P}\right)\qquad((p,m)=1).$$

于是由伽罗华理论，$(\mathbf{Z}/m\mathbf{Z})^\times/(\pm 1)\cong I_m(Q)/P_m(Q)$ 的所有商群(即 $I_m(Q)/H$，其中 H 是 $I_m(Q)$ 和 $P_m(Q)$ 的中间群)便给出 $Q(\zeta_m+\bar{\zeta}_m)$ 的所有子域. 根据 Kronecker-Weber 定理可知由此得到 Q 的所有实阿贝尔扩张.

设 $M=m\infty$，则 $I_M(Q)=I_m(Q)$，而

$$P_M(Q)=\left\{\left(\frac{a}{b}\right):a\text{ 和 }b\text{ 为互素的整数},\frac{a}{b}>0,\frac{a}{b}\equiv 1(\mathrm{mod}m)\right\}.$$

这时

$$C_M(Q)=\frac{I_M(Q)}{P_M(Q)}\cong(\mathbf{Z}/m\mathbf{Z})^\times.$$

于是在正则映射 $p\mapsto\left(\frac{Q(\zeta_m)/Q}{p}\right)$ 之下，$C_M(Q)$ 同构于分圆域 $Q(\zeta_m)$ 的伽罗华群. 所以 $C_M(Q)$ 的所有商群 $I_M(Q)/H\quad(P_M(Q)\subseteq H\subseteq I_M(Q))$ 给出分圆域的全部子域，由 Kronecher-Weber 定理，这给

出 Q 的所有阿贝尔扩张. 如果只用 $M=m$(m 过所有正整数),则只给出全体实阿贝尔域,这就是为什么还要把实位考虑进来的原因.

用这个例子为样板,高木得到关于数域阿贝尔扩张的一般结果.

定理 1 (1) 设 L/K 是数域的(有限)阿贝尔扩张. 则存在 K 的惟一的一个除子 f(叫作扩张 L/K 的**导子**),满足如下两个条件:

(i) K 的每个位 P 在 L 中分歧,当且仅当 P 是 f 的因子(有限位 P 即是 O_K 的素理想,而无限位 P(实位)在 L 中分歧,是指 P 到 L 的扩充为 L 的复位,即对应 L 到 $\boldsymbol{C}$ 的复嵌入).

(ii) 若 M 是 K 的除子,并且 $f|M$,则存在中间群 H,$P_M(K)\subseteq H\subseteq I_M(K)$,使得 $I_M(K)/H$ 正则同构于 $\mathrm{Gal}(L/K)$. 事实上 H 是由 L/K 和 M 所决定的:

$$H=P_M(K)\cdot N_{L/K}(I_M(L)),$$

其中 $I_M(L)$ 表示与 M_0 互素的 L 中分式理想构成的群,而对 L 的每个分式理想 A,$N_{L/K}(A)$ 表示 A 的所有共轭理想之积: $N_{L/K}(A)=\prod_\sigma \sigma(A)$, 其中 σ 过 $\mathrm{Gal}(L/K)$ 所有自同构. $N_{L/K}(A)$ 为 K 的分式理想. 而 H 是 $I_M(K)$ 的两个子群 $P_M(K)$ 和 $N_{L/K}(I_M(L))$ 的乘积,从而为 $I_M(K)$ 和 $P_M(K)$ 的中间群.

$I_M(K)/M$ 和 $\mathrm{Gal}(L/K)$ 的正则同构由 Artin 按如下方式构作出来:群 $I_M(K)$ 是由与 K 中整理想 M_0 互素的所有素理想 P 生成的. 由定理知 P 在 L 中不分歧(因为 $f|M$,从而 $p\nmid f$). 因此有 Frobenius 自同构 $\left(\frac{L/K}{P}\right)$. 由积性诱导出群同态

$$\varphi: I_M(K)\to\mathrm{Gal}(L/K),P\mapsto\left(\frac{L/K}{P}\right)\quad ((P,M_0)=1).$$

可以证明这是群的满同态,并且 $P_M(K)\subseteq\ker(\varphi)$. 进而可算出 $\ker(\varphi)=H=P_M(K)\cdot N_{L/K}(I_M(L))$(这些事实的证明都是很不平

凡的). 由此得到正则同构 $I_M(K)/H\cong\mathrm{Gal}(L/K)$, 并且 $P_M(K)\subseteq H\subseteq I_M(K)$. 这个同构叫作 Artin 映射.

(2) 反过来,给了 K 的一个除子 M 和中间群 H, $P_M(K)\subseteq H\subseteq I_M(K)$,则存在 K 的惟一有限阿贝尔扩域 L,使得

(i) 若 O_K 中(非零)素理想 P 在 L 中分歧,则 $P|M$.

(ii) 对于 $H=P_M(K)N_{L/K}(I_M(L))$,则 Artin 映射给出同构 $I_M(K)/H\cong\mathrm{Gal}(L/K)$.

(3) 设 L_1 和 L_2 均为 K 的有限阿贝尔扩域,导子分别为 f_1 和 f_2. 对于 K 的每个除子 M, $f_1|M$, $f_2|M$,则由(1)中方法得到

$$H_i=P_M(K)N_{L/K}(I_M(L_i))\quad(i=1,2).$$

我们有:$H_1\subseteq H_2$ 当且仅当 $L_1\supseteq L_2$.

这就是类域论的最基本结果. 简单说来,K 的所有有限阿贝尔扩域和 K 的所有模 M 广义类群的商群之间建立了一一对应,并且对应可通过 Artin 映射给出. (2)中的 M 叫作阿贝尔扩张 L/K 的模,最小可能的模 f 就是扩张 L/K 的导子. L 叫作(模 M)关于 H 的类域.

特别取 $M=1$,则 $I_1(K)=I(K)$, $P_1(K)=P(K)$. 则对于每个中间群 H, $P(K)\subseteq H\subseteq I(K)$, 则 K 有惟一的有限阿贝尔扩域 L,使得 $\mathrm{Gal}(L/K)\cong I(K)/H$,并且 O_K 的所有素理想(以及 K 的所有实位)在 L 中均不分歧. 如果取 $H=P(K)$ 为最小可能,由定理由(3)给出的反序对应,可知 $H=P(K)$ 对应 K 的最大不分歧阿贝尔扩域,这个域表示成 H_K,叫作 K 的希尔伯特类域. 于是 $\mathrm{Gal}(H_K/K)$ 正则同构于 K 的通常理想类群 $C(K)=I(K)/P(K)$. 特别地,$[H_K:K]=|C(K)|=h(K)$ 为 K 的理想类数. 如果我们能得到 K 的某个不分歧阿贝尔扩域 L,则 L 为 H_K 的子域,从而 $\mathrm{Gal}(L/K)$ 为 $\mathrm{Gal}(H_K/K)\cong C(K)$ 的商群,而 $[L:K]$ 为 $[H_K:K]=h(K)$ 的因子. 由此可研究 K 的理想类群结构和类数 $h(K)$ 的整除性.

例 考虑虚二次域 $K=\boldsymbol{Q}(\sqrt{-5})$. 域 $L=\boldsymbol{Q}(\sqrt{-5},\sqrt{-1})$ 是 K

的二次扩张，从而必然是阿贝尔扩张. 容易证明 O_K 的每个素理想在 L 中均不分歧，而 K 没有实位($r_1=0$)，所以 L/K 是不分歧阿贝尔扩张. 于是 $2=[L:K]$为 K 的类数 $h(K)$的因子. 事实上 $h(K)=2=[L:K]$，所以 L 就是 K 的最大不分歧阿贝尔扩张，即 $\boldsymbol{Q}(\sqrt{-5},\sqrt{-1})=K(\sqrt{-1})$是 $K=\boldsymbol{Q}(\sqrt{-5})$的希尔伯特类域. 类似方法可证明 $\boldsymbol{Q}(\sqrt{2},\sqrt{5})$是 $K=\boldsymbol{Q}(\sqrt{5})$的类域，$h(K)=2$.

对于每个阿贝尔数域 L(即 L/Q 为阿贝尔扩张)，L 的导子有形式 f 或 $f\infty$，则 f 是使 $L\subseteq \boldsymbol{Q}(\zeta_f)$的最小正整数，即 $\boldsymbol{Q}(\zeta_f)$是包括 L 的最小分圆域. 我们有时也把 f 叫作 L 的导子. 当 $L=\boldsymbol{Q}(\sqrt{d})$为二次域时，用分歧理论不难证明 L 的导子就是判别式的绝对值 $|d(L)|$(当 d 为无平方因子整数时，$d(L)$为 d 或 $4d$). 即 $\boldsymbol{Q}(\zeta_{|d(L)|})$是包含二次域 L 的最小分圆域.

类域论还部分地解决了希尔伯特另一个问题(第 9 问题)，我们在第四章再介绍.

参 考 文 献

(类域论有许多著作，[1]是最好的一本入门书)

[1] J. Neukirch, Klassenkorpertheorie, 1969. 英版本：Class Field Theory, Springer-Verlag, 1986.

[2] H. Hasse, History of class field theory, In Cassels-Fröhlich, Algebraic Number Theory, Academic Press, 1967, p. 266～279.

§3.2 局部域和局部-整体原则(20世纪30年代)

希尔伯特第11问题是研究任意数域上的多变量二次型方程的解.代数数域 K 上多元二次型为

$$f(x_1,\cdots,x_k)=\sum_{i,j=1}^{k}a_{ij}x_ix_j \quad (a_{ij}\in K).$$

希尔伯特问:

(1) 对 K 中每个元素 $\alpha\in K$,方程 $f(x_1,\cdots,x_k)=\alpha$ 是否在 K 中有解 $(x_1,\cdots,x_k)(x_i\in K)$?若 $\alpha\neq 0$ 而此方程在 K 中有解,我们称 f(在 K 中)表示 α.若 $\alpha=0$,则方程显然有零解 $(x_1,\cdots,x_k)=(0,\cdots,0)$.如果 $f(x_1,\cdots,x_k)=0$ 在 K 中有非零解,我们称 f(在 K 中)表示0.于是,这个问题还可以说成:K 上的二次型 $f(x_1,\cdots,x_k)$ 可以表示 K 中哪些元素?

这个问题由德国数论学家哈瑟(Hasse,1898—1981)于20世纪30年代给出了一个相当满意的答案(至少是一个漂亮的理论结果).他由此创造了数论(乃至代数几何等学科)中一种新的研究方法,叫局部—整体原则.

希尔伯特的这个问题还有另一种提法:

(2) 如果二次型 $f(x_1,\cdots,x_k)$ 的系数 a_{ij} 均属于 K 的整数环 O_K,那么 $f(x_1,\cdots,x_k)$ 在 O_K 上可表示哪些整数?即对 $\alpha\in O_K$,方程 $f(x_1,\cdots,x_k)=\alpha$ 是否有(非零)整解 $(x_1,\cdots,x_k)$,其中 $x_i\in O_K$?

例如对于 $K=\boldsymbol{Q}$,从古代东方和古希腊就开始研究的二平方和问题 $x^2+y^2=n$,Pell 方程 $x^2-dy^2=1$,勾股方程 $x^2+y^2-z^2=0$ 的整数解问题,甚至高斯二元二次型的一般性研究,都是问题(2)

的特殊情形. 整解问题比数域上的问题(1)要困难,对此问题的研究促成了另一个数论分支—模形式理论的建立,我们将在§3.4中介绍.

现在介绍哈瑟对问题(1)给出的解法.

为了简单起见,我们用 $K=\boldsymbol{Q}$ 来说明哈瑟的思想,然后叙述对任意代数数域 K 的哈瑟结果.

考虑方程

$$f(x_1,\cdots,x_k)=\sum_{i,j=1}^{n}a_{ij}x_ix_j=\alpha \quad (a_{ij},\alpha\in\boldsymbol{Q}). \qquad (1)$$

研究此方程是否有(非零)有理数解,今后我们不妨假定二次型 f 是非退化的,即方阵 (a_{ij}) 的行列式不为 0(因若 f 退化,由线性代数的二次型理论知方程(1)可等价于一个变量更少的非退化二次型方程).

如果方程(1)在 $\boldsymbol{Q}$ 中有解,则它也是实数域 $\boldsymbol{R}$ 中的解. 方程(1)是否有实数解则容易判断:对称方阵 (a_{ij})(不妨设 $a_{ij}=a_{ji}$)在实数域上可以对角化,使得对角元素为 ±1. 如果对角元素均为 1(即 (a_{ij}) 为正定的),则方程(1)有实数解当且仅当 $\alpha>0$. 若对角元素均为 -1(即 (a_{ij}) 负定),则(1)有实数解当且仅当 $\alpha<0$. 最后若对角线既有 1 也有 -1(即 (a_{ij}) 是不定型的),则方程(1)对任何 α 均有实数解. 实数解的判别所以这样简单,根本原因是:实数域 $\boldsymbol{R}$ 为 $\boldsymbol{Q}$ 对于通常绝对值给出的拓朴的完备化距离空间,所以在 $\boldsymbol{R}$ 中可以取极限,有连续性概念,甚至可以使用解析工具(微积分),人们要问:有理数域 $\boldsymbol{Q}$ 是否还有其他的距离概念(即衡量有理数大小和彼此距离是否有其他标准),并且 $\boldsymbol{Q}$ 对于这个拓朴作成完备化域 $\boldsymbol{Q}'$,使得 $\boldsymbol{Q}'$ 中有解析工具,从而容易判别方程(1)在 $\boldsymbol{Q}'$ 中是否有解?进而,由于 $\boldsymbol{Q}'$ 是 $\boldsymbol{Q}$ 的扩域,所以方程在 $\boldsymbol{Q}$ 中的解自然也是在 $\boldsymbol{Q}'$ 中的解. 反过来,如果我们构作了许多的 $\boldsymbol{Q}'$,如果方程(1)在所有的 $\boldsymbol{Q}'$ 中均有解,是否能够推出方程(1)在 $\boldsymbol{Q}$ 中有解?

哈瑟在30年代的这种思考是有基础的，因为在1900年前后，另一个德国人Hensel（1861—1941，库默尔的学生）就已经迈出了第一步（我们曾经说过，库默尔在研究分圆域中单位性质时已具有了Hensel p-adic思想的原型）. Hensel发现了比较有理数大小的无穷多种不同的距离标准，称作为**赋值**. 从泛函分析中我们知道，域F中的一个距离是一个映射$\varphi: F\to \boldsymbol{R}$，并且满足如下三个条件：对$a,b\in F$，

（1）$\varphi(a)\geqslant 0$，并且$\varphi(a)=0$当且仅当$a=0$；

（2）$\varphi(ab)=\varphi(a)\varphi(b)$；

（3）（三角形不等式）$\varphi(a+b)\leqslant\varphi(a)+\varphi(b)$.

有理数域对于通常的绝对值$\varphi(a)=|a|$满足这三个条件，并且$\boldsymbol{Q}$对于由它决定的拓朴是豪斯道夫拓朴空间，其完备化就是实数域$\boldsymbol{R}$. 现在设p为素数，对每个非零整数a，令$p^l\,\|\,a$（这表示l为非负整数，满足$p^l|a, p^{l+1}\nmid a$. 换句话说，l是a素因子分解式中p的指数）. Hensel定义$v_p(a)=l$. 并且令$v_p(0)=+\infty$. 最后，对每个非零有理数$\alpha=\frac{a}{b}(a,b\in\boldsymbol{Z}, ab\neq 0)$，定义$v_p(\alpha)=v_p(a)-v_p(b)$. 不难验证：$v_p$是从$\boldsymbol{Q}$到$\boldsymbol{Z}\cup\{+\infty\}$中的映射，满足

（1′）$v_p(\alpha)=+\infty$当且仅当$\alpha=0$；

（2′）$v_p(\alpha\beta)=v_p(\alpha)+v_p(\beta)$；

（3′）$v_p(\alpha+\beta)\geqslant\min(v_p(\alpha), v_p(\beta))$.

v_p叫作$\boldsymbol{Q}$上的p-adic **指数赋值**. 如果取一个实数$\gamma, 0<\gamma<1$. 令$|\alpha|_p=\gamma^{v_p(\alpha)}$（对$\alpha\in\boldsymbol{Q}$），则$|\ |_p$是从$\boldsymbol{Q}$到$\boldsymbol{R}$中的映射（由于$0<\gamma<1$，我们可认为$\gamma^{+\infty}=0$，即$|0|_p=0$），并且性质（1′），（2′）和（3′）分别变成如下三个性质：对$\alpha,\beta\in\boldsymbol{Q}$，

（1）$|\alpha|_p\geqslant 0$，并且$|\alpha|_p=0$当且仅当$\alpha=0$；

（2）$|\alpha\beta|_p=|\alpha|_p\cdot|\beta|_p$；

（3）$|\alpha+\beta|_p\leqslant\max(|\alpha|_p, |\beta|_p)(\leqslant|\alpha|_p+|\beta|_p)$.

这就表明$|\ |_p$是$\boldsymbol{Q}$上的一种距离，叫作$\boldsymbol{Q}$的p-adic赋值. 而a和b

的 p-adic 距离为$|a-b|_p$. 由于$0<\gamma<1$,所以对于 p-adic 赋值来说,一个整数 a 很小,即$|a|_p=\gamma^{v_p(a)}$很小,是指 $v_p(a)$很大,即 a 被 p 的很高次幂除尽. 例如 $v_3(\pm 3)=1, v_3(\pm 5)=0$,于是$|\pm 3|_3=\gamma <1=\gamma^0=|\pm 5|_3$. 但是$|\pm 3|_5=1>\gamma=|\pm 5|_5$. 对于不同的素数 p 和 q,p-adic 赋值和 q-adic 赋值给出完全不同的大小和距离概念,即给出 $\boldsymbol{Q}$ 上不同的拓朴. 对于同一个 p-adic 赋值,取不同的 γ 值,则给出 $\boldsymbol{Q}$ 上同一种拓朴结构,我们把不同取值 γ 的所有 p-adic 赋值叫作彼此等价的,统称为 $\boldsymbol{Q}$ 的一个**素除子**,并且就表示成 p. 它们均叫**非阿**(基米德)**素除子**,或**有限素除子**,它们不仅满足三角形不等式,而且满足比它更强的不等式

$$|\alpha+\beta|_p\leqslant\max(|\alpha|_p,|\beta|_p), \tag{2}$$

所以比通常绝对值有很多特殊的性质(非阿特性). 比如说:

若$|\alpha|_p\neq|\beta|_p$,则$|\alpha\pm\beta|_p=\max(|\alpha|_p,|\beta|_b)$.

也就是说,对于非阿距离,$\boldsymbol{Q}$ 上任何"三角形"一定是等腰三角形(若$|a-b|_p\neq|a-c|_p$,则第三边长$|b-c|_p$ 必为$|a-b|_p$ 和 $|a-c|_p$的最大者). 以 a 为圆心 r 为半径的 p-adic"圆"是满足$|a-x|_p\leqslant r$ 的所有有理数 x. 圆内任意一点都可作此圆的圆心,即若$|a-b|_p<r$,则这个圆也可表成$|b-x|_p\leqslant r$.

通常的绝对值$|\ |$(的等价类)叫作**阿基米德赋值**,或**无限赋值**,表示成∞. Hensel 证明有理数域上由此给出了全部可能的"距离",即 $p=2,3,5,7,\cdots$以及 $p=\infty$为 $\boldsymbol{Q}$ 上的全部素除子. 以它们为基的自由阿贝尔群叫作 $\boldsymbol{Q}$ 的**除子群**,每个元素叫作一个除子(这就是上节所述的除子概念).

$|\ |_p$ 叫 p-adic(p 进)赋值,是因为它与有理数的 p-adic(p 进)展开有关. 每个正整数 a 可以 p 进展成

$$a=c_lp^l+c_{l+1}p^{l+1}+\cdots+c_np^n$$

其中 c_i 为整数,$0\leqslant c_i\leqslant p-1$,并且 $c_l\neq 0$,于是 $p^l\,\|\,a$,即 $v_p(a)=l$, $|a|_p=\gamma^l$(今后我们尽量采用 p-adic 指数赋值 v_p,因为它的定义与 γ 的选取无关).

$\boldsymbol{Q}$ 对 p-adic 拓朴是豪斯道夫拓朴空间，其完备化是域，叫作 p-adic 数域，表示成 $\boldsymbol{Q}_p$，而 $\boldsymbol{Q}$ 对通常绝对值的完备化为实数域 $\boldsymbol{R}$，它也表成 $\boldsymbol{Q}_\infty$. $\boldsymbol{Q}_p$ 中每个数 α 叫 p-adic 数，当 $\alpha \neq 0$ 时它可以表示成无穷 p-adic 展开

$$\alpha = c_l p^l + c_{l+1} p^{l+1} + \cdots + c_n p^n + \cdots = \sum_{n=l}^{\infty} c_n p^n,$$

其中 c_i 均为整数，$0 \leqslant c_i \leqslant p-1$，而 $c_l \geqslant 1$（于是 $v_p(\alpha)=l$）. 如果 p-adic 展开是周期的，即存在整数 t 和 $s(s\geqslant 1)$，使得当 $n \geqslant t$ 时，$c_{n+s} = c_n$. 则 α 为有理数（就像在 $\boldsymbol{R}=\boldsymbol{Q}_\infty$ 中有理数都是循环小数一样）. 于是，具有非周期性 p-adic 展开的 p-adic 数就是由 $\boldsymbol{Q}$ 扩大出来的（对 $\boldsymbol{R}=\boldsymbol{Q}_\infty$，非循环小数即是无理数）.

由于 $\boldsymbol{Q}_p$ 是完备拓朴空间，在其中（对 p-adic 拓朴）有连续性概念，可以取极限，加法和乘法运算都是 p-adic 连续函数，从而 $\boldsymbol{Q}_p$ 是拓朴域，非阿拓朴域 $\boldsymbol{Q}_p(p \neq \infty)$ 有简单的拓朴和代数结构，令

$$\boldsymbol{Z}_p = \{\alpha \in \boldsymbol{Q}_p \mid |\alpha|_p \leqslant 1\} = \{\alpha \in \boldsymbol{Q}_p \mid v_p(\alpha) \geqslant 0\},$$

由非阿性质(2)式，可知 $\boldsymbol{Z}_p$ 是 环，叫 p-adic **整数环**，$\boldsymbol{Z}_p$ 中元素叫 p-adic 整数. $\boldsymbol{Z}$ 为 $\boldsymbol{Z}_p$ 的子环，并且 $\boldsymbol{Z}_p$ 是 $\boldsymbol{Z}$ 在 $\boldsymbol{Q}_p$ 中的紧闭包，$\boldsymbol{Z}_p$ 的全部非零理想为

$$(p^n) = p^n \boldsymbol{Z}_p \quad (n=0,1,2,3,\cdots),$$

其中 $(p^n) = p^n\boldsymbol{Z}_p = \{\alpha \in \boldsymbol{Z}_p \mid v_p(\alpha) \geqslant n\} = \{\alpha \in \boldsymbol{Z}_p \mid v_p(\alpha) > n-1\}$ 可知 (p^n) 是又开又闭的集合. 这些理想形成 0 的基本邻域系，所以 $\boldsymbol{Q}_p$ 对于 p-adic 拓朴是全不连通的拓朴空间. 环 $\boldsymbol{Z}_p$ 的可逆元素全体形成单位群表示成 U_p. U_p 中元素 α（即 $v_p(\alpha)=0$ 或 $|\alpha|_p=1$）叫 p-adic 单位. 群 U_p 也有简单的代数结构：

$$U_p = W_p \times (1+p\boldsymbol{Z}_p) \text{（直积）},$$

其中 W_p 是 $p-1$ 次单位根组成的 p-1 阶循环群，子群 $1+p\boldsymbol{Z}_p$ 还可细分出一些更小的子群

$$U^{(n)} = 1 + p^n \boldsymbol{Z}_p \quad (n=1,2,3,\cdots).$$

它们也是又开又闭的子群，并且构成拓朴群 $\boldsymbol{Q}_p^*$ 中 1 的基本邻域

系，因此乘法群 $\boldsymbol{Q}_p^x=\boldsymbol{Q}_p-\{0\}$ 也是全不连通的拓朴空间.

我们可以用下述方式证明 U_p 中存在 $p-1$ 次本原单位根：对每个整数 $a(1\leqslant a\leqslant p-1)$，初等数论有 $a^{\varphi(p^n)}\equiv1(\bmod\ p^n)$(欧拉定理)，由于 $\varphi(p^n)=p^n-p^{n-1}$，所以 $a^{p^n}\equiv a^{p^{n-1}}(\bmod\ p^n)$ 于是 $v_p(a^{p^n}-a^{p^{n-1}})\geqslant n$，即 $a^{p^n}(n=0,1,2,\cdots)$ 是柯西序列(由于非阿性质(2)，$\boldsymbol{Q}_p$ 中序列 $\{c_0,c_1,\cdots,c_n,\cdots\}$ 为柯西序列，当且仅当 $|c_n-c_{n-1}|_p\to0$，即 $v_p(c_n-c_{n-1})\to+\infty$(当 $n\to\infty$ 时)，这比实拓朴要简单)，从而有 p-adic 极限 $\alpha_a=\lim\limits_{n\to\infty}a^{p^n}$. 由于 $\boldsymbol{Q}_p$ 对 p-adic 拓朴完备，于是 $\alpha_a\in\boldsymbol{Q}_p$，但是 $v_p(a^{p^n})=p^nv_p(a)=0$. 由赋值连续性，$v_p(\alpha_a)=0$. 即 $\alpha_a\in U_p$. 由于 $v_p(a^{p^n(p-1)}-1)\geqslant n+1$(即 $a^{p^n(p-1)}\equiv1(\bmod\ p^{n+1})$)取极限可知 $\alpha_a^{p-1}=1$，即 $\alpha_a(1\leqslant a\leqslant p-1)$ 都是 $p-1$ 次单位根. 易知这 $p-1$ 个 p-adic 单位彼此不同(因为 $\alpha_a\equiv a(\bmod\ p)$)，所以它们当中必有 $p-1$ 次本原单位根.

我们用这个例子向大家展示 p-adic 分析和实分析相比有很不相同的特点. 事实上，在 $\boldsymbol{Q}_p$ 上(p 为素数)解代数方程与在 $\boldsymbol{R}=\boldsymbol{Q}_\infty$ 上相比，有一系列风格完全不同的方法(Hensel 引理，牛顿折线法等等). 特别地，在 $\boldsymbol{Q}_p$ 中解二次型方程(1)，哈瑟采用希尔伯特在《数论报告》中引入的希尔伯特符号，可以通过简单的计算，相当容易地判别方程(1)在 $\boldsymbol{Q}_p$ 中何时有(非零)解. 由于篇幅所限，具体细节从略.

完全类似地，对每个代数数域 K，取 P 为 O_K 的一个非零素理想. 对每个 $0\neq\alpha\in O_K$，主理想 αO_K 有素理想分解式 $\alpha O_K=\prod\limits_P P^{a_P}$(有限乘积)，$a_P\geqslant0$，定义 $v_P(\alpha)=a_P$，$v_P(0)=+\infty$. 对每个 $0\neq\alpha=\dfrac{a}{b}\in K^x$，$(a,b\in O_K)$，令 $v_P(\alpha)=v_P(a)-v_P(b)$. 则 v_P 为从 K 到 $\boldsymbol{Z}\cup\{+\infty\}$ 的映射，并且满足前面的条件(1′)～(3′). 仍取 $0<\gamma<1$，对 $\alpha\in K$ 定义 $|\alpha|_P=r^{v_P(\alpha)}$，则 $|\ |_P$ 是非阿赋值，v_P 和 $|\ |_P$ 分别叫 K

的 P-adic 指数赋值和 P-adic 赋值. 对于 O_K 的不同素理想 P 和 Q，K 中 P-adic 赋值和 P-adic 赋值给出不同的拓朴结构，K 对 P-adic 赋值的完备化表示成 K_P，于是有无穷多个非阿完备赋值域 K_P(P 过 O_K 的所有非零素理想)，集合

$$O_{K,P}=\{\alpha\in K\,|\,v_P(\alpha)\geqslant 0\}=\{\alpha\in K\,|\,v_P(\alpha)>-1\}$$

是又开又闭的紧环，叫作 K 的 P-adic 整数环，它的所有非零理想为 P^n 到 $O_{K,P}$的扩充 $P^nO_{K,P}(n=0,1,2,\cdots)$存在元素 $\Pi\in O_{K,P}$，使得 $v_P(\Pi)=1$，称 Π 为 P-adic **素元**，则 $P^nO_{K,P}=(\Pi^n)=\Pi^nO_{K,P}$. 于是 $O_{K,P}$是主理想整环，并且(Π)是 $O_{K,P}$的惟一极大理想，从而 $O_{K,P}$ 是局部环(对于 O_K，只有 K 的理想类数 $h(k)$为 1 时才是主理想整环，所以 $O_{K,P}$比 O_K 的环结构要简单许多)，我们也把 K_P 叫作**局部域**. 或叫作 K 在 P 处的**局部化**. 除了这些非阿局部域之外，K 还有一些对阿基米德赋值的完备化域. 设 $\sigma_i(1\leqslant i\leqslant n)$是 K 到 $\boldsymbol{C}$ 的 n 个嵌入，$n=[K:\boldsymbol{Q}]$. 其中前 r_1 个是实嵌入，而后 r_2 对为复嵌入，$r_1+2r_2=n$，$\sigma_{r_1+j}=\bar{\sigma}_{r_1+r_2+j}(1\leqslant j\leqslant r_2)$，记$|\ |$为通常的绝对值，则对每个 $i(1\leqslant i\leqslant n)$，定义

$$|\alpha|_{\infty,i}=|\sigma_i(\alpha)|\qquad(\alpha\in K),$$

则$|\ |_{\infty,i}$均是 K 的阿基米德赋值. 由于对每个 $\alpha\in K$，

$$|\alpha|_{\infty,r_1+j}=|\sigma_{r_1+j}(\alpha)|=|\bar{\sigma}_{r_1+r_2+j}(\alpha)|=|\sigma_{r_1+r_2+j}(\alpha)|=$$

$|\alpha|_{\infty,r_1+r_2+j}$，所以$|\ |_{\infty,r_1+j}$和$|\ |_{\infty,r_1+r_2+j}$是同样的赋值. 可以证明：$|\ |_{\infty,i}(1\leqslant i\leqslant r_{1+}r_2)$给出 K 的 $r_{1+}r_2$ 个不同的拓朴. 并且 K 对$|\ |_{\infty,i}$的完备化前 r_1 个为实数域 $\boldsymbol{R}$，后 r_2 个为复数域 $\boldsymbol{C}$，所以 K 有 r_1 个实位和 r_2 个复位. $\boldsymbol{R}$ 和 $\boldsymbol{C}$ 也叫 K 对实位和复位的局部化，表示成 $K_{\infty,i}=\boldsymbol{R}(1\leqslant i\leqslant r_1)$，$K_{\infty,i}=\boldsymbol{C}(r_1+1\leqslant j\leqslant r_1+r_2)$.

如果 K 上的二次型方程(1)在 K 中有非零解，则对每个位 P(包括阿基米德位$\infty_i(1\leqslant i\leqslant r_1+r_2)$)，它也是局部域 K_P 中的非零解(因为 $K_P\geqslant K$)，哈瑟证明了反过来也成立.

定理 代数数域 K 上的二次型方程(1)在 K 中有(非零)解当且仅当方程(1)在每个局部域 K_P(包括 K 的所有阿基米德位 $P=\infty$)中均有(非零)解.

可以证明:对每个方程(1),它在几乎所有局部域 K_P 中均有解(几乎所有是指除了有限个 K_P 之外),所以即使局部域是无限多个,我们只需考查方程(1)在有限个局部域中是否有解,而在每个局部域中,可以用希尔伯特符号的简单计算就可判别方程(1)是否有解.这就给出判别方程(1)在域 K 中是否有解的一种好方法.

比如说,若 $f(x_1,\cdots,x_k)$ 是系数属于 K 的非退化二次型,可以证明,当 $k\geqslant 4$ 时,对每个 $\alpha\in K^x$,方程 $f(x_1,\cdots,x_k)=\alpha$ 在每个非阿局部域中有解,由于它显然在 $\boldsymbol{C}$ 中有解,所以若 K 不是实的数域(即 K 不为 $\boldsymbol{R}$ 的子域),则当 $k\geqslant 4$ 时,$f(x_1,\cdots,x_k)=\alpha$ 在 K 中必有解,即 $f(x_1,\cdots,x_k)$ 可表示 K 中每个非零元素,而当 $K\subseteq\boldsymbol{R}$ 时,$f(x_1,\cdots,x_k)=\alpha(\alpha\in K^x)$在 K 中有解当且仅当此方程有实解.

类似地,若 $k\geqslant 5$,则 $f(x_1,\cdots,x_k)$在每个非阿局部域 K_P 中可表示 0(即 $f(x_1,\cdots,x_k)=0$ 在 K_P 中有非零解),所以若 K 不是实的,则当 $k\geqslant 5$ 时 $f(x_1,\cdots,x_k)=0$ 在 K 中必有非零解.若 $K\subseteq\boldsymbol{R}$,则当 $k\geqslant 5$ 时,$f(x_1,\cdots,x_k)=0$ 在 K 中有非零解当且仅当 f 是不定的二次型.

哈瑟的这种方法在后来的数论研究中得到普遍的应用.如果数域 K(也叫整体域)上的一个性质是局部性质,则到局部域 K_P 上去考虑,问题常常变得简单.或者把在所有局部域 K_P 上的信息收集起来,试图得到整体域中的某些结论,这种思想后人称之为哈瑟的局部-整体原则.关于二次型的上述结果是使用这种原则的最漂亮结果,对大多数情形还会更复杂些.例如对有理数域上的方程

$$f(x)=(x^2-13)(x^2-17)(x^2-221)=0.$$

这个方程显然有实数解,因为$\sqrt{13}$,$\sqrt{17}$和$\sqrt{221}$均为实数.由

Hensel 定理可以证明：对于 $a\in \mathbf{Z}$，

（1）若 $a\equiv 1 \pmod 8$，则方程 $x^2=a$ 在 $\boldsymbol{Q}_2$ 中有解.

（2）对于奇素数 p，则当 $\left(\frac{a}{p}\right)=1$ 时，方程 $x^2-a=0$ 在 $\boldsymbol{Q}_p$ 中有解.

由于 $11\equiv 1 \pmod 8$，所以以上方程在 $\boldsymbol{Q}_2$ 中有解. 由 $\left(\frac{17}{13}\right)=\left(\frac{13}{17}\right)=1$，可知上方程在 $\boldsymbol{Q}_{13}$ 和 $\boldsymbol{Q}_{17}$ 中均有解. 当 p 为 13 和 17 之外的奇素数时，若 $\left(\frac{13}{p}\right)=1$ 或 $\left(\frac{17}{p}\right)=1$，则 $x^2-13=0$ 或 $x^2-17=0$ 在 $\boldsymbol{Q}_p$ 中有解. 若 $\left(\frac{13}{p}\right)=\left(\frac{17}{p}\right)=-1$，则 $\left(\frac{221}{p}\right)=\left(\frac{13}{p}\right)\left(\frac{17}{p}\right)=1$，从而 $x^2-221=0$ 在 $\boldsymbol{Q}_p$ 中有解. 于是上方程在 $\boldsymbol{Q}_p$ 中必有解. 这表明上方程在 $\boldsymbol{Q}$ 的每个局部域中均有解，但是它在 $\boldsymbol{Q}$ 中显然无解，因为 $\sqrt{13}$，$\sqrt{17}$ 和 $\sqrt{221}$ 都不是有理数.

关于局部域以及局部域与整体域的各种联系，已有相当完整的理论，也发现了用局部域 K 的内部性质刻划 K 的所有阿贝尔扩张的理论，叫局部类域论（它比研究数域阿贝尔扩张的整体类域论要简单许多）. 这些局部域理论对后来数论研究起很重要的作用.

参 考 文 献

[1] T. Y. Lam, The Algebraic Theory of Quadratic Forms. 1973.

[2] J. P. Serre, Cours d' arithmétique, 1970（英文本：A Course

in Arithmétique, 1973. 中译本：数论教程，上海科学技术出版社，1980)([2]中讲述 $\boldsymbol{Q}$ 上二次型的局部-整体原则，[1]中讲述任何数域的情形).

[3] J.-P. Serre, Local Felds, GTM67, Springer-Verlag(讲述局部域和局部类域论).

[4] K. Iwasawa, Local Class Field Theory, Oxford Univ. Press, 1986(有中译本).

§3.3 有限域上函数域的算术(20世纪20～40年代)

Hensel 的赋值理论在数论中发展成局部域的理论、局部化方法和局部-整体原则. 后来又发展出 p-adic 分析作为研究数论的重要工具. 在本节中我们要介绍赋值理论对几何学的重要影响，以及代数几何学在 20 世纪 40 年代如何与代数数论交织在一起，产生出算术几何这个交叉性分支.

在古代，数论是依属于几何的，数用对几何图形的测量体现出来. 无理数的发现也是古希腊几何学的产物(边长为 1 的正方形对角线长度). 到了 18 世纪，随着对整数研究的不断深入，数论才形成学科而与几何学相对独立. 19 世纪代数数论的发展，使数论更加独立而脱离几何的轨道，但是赋值论又把它们联系在一起，而这次是数论得到很大的发展，几何向数论靠近.

设 $f(x,y)\in\boldsymbol{R}[x,y]$，即 $f(x,y)$ 是实系数多项式，方程 $f(x,y)=0$ 是坐标平面上的一条曲线，这条曲线所有点组成 $\boldsymbol{R}^2$ 的一个子集合，就是 $f(x,y)=0$ 的全部实数解，这是几何学研究对象，如果 $f(x,y)\in\boldsymbol{Z}[x,y]$，而考虑 $f(x,y)=0$ 的全体整数解，就是数

论.按照德国数学家克莱因(F. Klein)的观点(Erlanger 纲领),几何学是研究某些对象在某个群作用下不变量的理论,代数几何学是研究域 F 上代数方程组的解的性质.设 $f_i(x_1,\cdots,x_k)\in F[x_1,\cdots,x_k]$ $(1\leqslant i\leqslant n)$,则方程组 $f_i(x_1,\cdots,x_k)=0(1\leqslant i\leqslant n)$ 在 F 中的公共解全体形成 F^k 的一个子集合,叫作代数簇.一维代数簇就是 F 上的代数曲线,这些代数簇是代数几何的研究对象,代数几何的作用群是代数簇之间的双有理变换群.

设 V_1 和 V_2 是域 E 上的两个代数簇,V_1 是由 E^n 中关于 $x_1,\cdots,x_n$ 的代数方程组(Ⅰ)定义的,V_2 是由 E^m 中关于 $y_1,\cdots,y_m$ 的代数方程组(Ⅱ)定义的.如果存在双方的有理变换

$$\begin{cases} x_1=\dfrac{f_1(y_1,\cdots,y_m)}{g_1(y_1,\cdots,y_m)}, \\ \vdots \\ x_n=\dfrac{f_n(y_1,\cdots,y_m)}{g_n(y_1,\cdots,y_m)}, \end{cases} \qquad \begin{cases} y_1=\dfrac{F_1(x_1,\cdots,x_n)}{G_1(x_1,\cdots,x_n)}, \\ \vdots \\ y_m=\dfrac{F_m(x_1,\cdots,x_n)}{G_m(x_1,\cdots,x_n)}. \end{cases}$$

其中

$$f_l(y_1,\cdots,y_m),g_l(y_1,\cdots,y_m)\in E[y_1,\cdots,y_m](1\leqslant l\leqslant n),$$
$$F_k(x_1,\cdots,x_n),G_k(x_1,\cdots,x_n)\in E[x_1,\cdots,x_n](1\leqslant k\leqslant m),$$

使得

(1) 若 $(y_1,\cdots,y_m)=(a_1,\cdots,a_m)$ 是代数簇 V_2 的一个点(即它是方程组(Ⅱ)在域 E 中的一个解),并且 $g_i(a_1,\cdots,a_m)\neq 0(1\leqslant i\leqslant n)$.把它代入第一个变换公式之后,$x_i=\dfrac{f_i(a_1,\cdots,a_m)}{g_i(a_i,\cdots,a_m)}(1\leqslant i\leqslant n)$ 是 V_1 的点(即为方程组(Ⅰ)在域 E 中的解).同样地,若 $(x_1,\cdots,x_n)=(b_1,\cdots,b_n)$ 是 V_1 的一个点,并且 $G_j(b_1,\cdots,b_n)\neq 0(1\leqslant j\leqslant m)$,则 $y_j=\dfrac{F_j(b_1,\cdots,b_n)}{G_j(b_1,\cdots,b_n)}(1\leqslant j\leqslant m)$ 是 V_2 上的点.

(2) 除了一个低维子代数簇(即除去使双有理变换中有理函数的分母为 0 的那些点)之外,上面变换给出互逆的映射.

这时称代数簇 V_1 和 V_2 是双有理等价的. 由定义可知,对于

双有理等价的两个代数簇，定义它们的两个方程组在域 E 中的解本质上通过双有理变换是一一对应的（所谓“本质上”是指除去使双有理变换中有理函数的分母为零的那些点）．换句话说，从一个方程组在 E 中的全部解本质上（通过有理变换）可得到另一个方程组的全部解．这就是研究代数簇双有理等价分类的原始动机．

代数几何的基本问题是对代数簇进行双有理等价分类：如何判别两个代数簇是否双有理等价？这就需要研究代数簇的双有理不变量和不变性．一个直观上易于接受的双有理不变量是代数簇的维数（但是要严格证明双有理等价的代数簇有相同的维数，甚至对维数作严格的数学定义都是不简单的），于是，曲线只能和曲线双有理等价．曲线的另一个重要的双有理不变量叫作亏格(genus)，是由黎曼发现的，它是非负整数．亏格为 0 的曲线即是直线和二次曲线，它们形成一个等价类．而对每个正整数 g，亏格为 g 的曲线仍分成无穷多等价类．高维代数簇的分类更加困难，至今也没有解决．

我们举例说明二次曲线在任何域 F 上都双有理等价于直线．设 V_1 是平面 F^2 上的单位圆周 $x^2+y^2=1$，V_2 是直线 F（以 t 为变量，没有方程）．过点 $A=(-1,0)$ 作斜率为 t 的直线

$$l: y=t(x+1).$$

这条直线和单位圆有两个交点：A 和 B．容易算出点 B 为 $(x,y)=\left(\frac{1-t^2}{1+t^2},\frac{2t}{1+t^2}\right)$．当斜率 t 过 F 中所有元素，则点 B 走过单位圆周上除了 A 之外的所有点．因此，单位圆除了点 A 之外，和 t-直线在双有理变换下一一对应，变换为

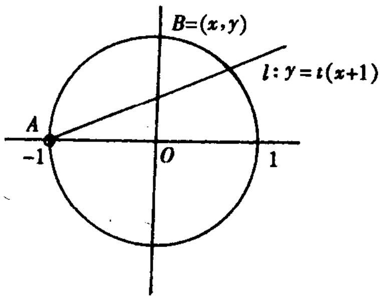


图 11

$$\begin{cases} x=\dfrac{1-t^2}{1+t^2}, \\ y=\dfrac{2t}{1+t^2}, \end{cases} \qquad t=\frac{y}{x+1}.$$

设 $C:f(X,Y)=0$ 是域 F 上一条代数曲线，其中 $f(X,Y)\in F[X,Y]$，曲线 C 的函数域为 $F(C)=F(x,y)$，其中 x 和 y 满足代数方程 $f(x,y)=0$. 于是 y 在有理函数域 $k=F(x)$ 上是代数的，而函数域 $F(C)=k(y)$ 是有理函数域 k 的有限扩域，这就象代数数域是有理数域 Q 的有限扩域一样. 19 世纪随着代数学(特别是域论)的发展，发现曲线 C_1 和曲线 C_2 在域 F 上双有理等价，本质上相当于它们的函数域 $F(C_1)$ 和 $F(C_2)$ 是 F 同构的(即同构将 F 上每个元素保持不变). 这就把几何问题代数化，集中到对函数论的研究. 由于函数域和数域的类似，19 世纪末期狄德金和 Kronecker 开始用代数数域的思想研究这些函数域.

最基本的数域和函数域分别是有理数域 Q 和有理函数域 $k=F(x)$，而一般的数域和函数域分别是它们的有限扩域. 有理函数域中的多项式环 $F[x]$ 类比于有理数域中的整数环 $\mathbf{Z}$. 它们都是主理想整环，从而都有惟一因子分解性(UFD). $F[x]$ 中的首 1 不可约多项式相当于 $\mathbf{Z}$ 中的素数. $F[x]$ 和 $\mathbf{Z}$ 的单位群分别为 $F^{\times}=F-\{0\}$ 和 $\{\pm 1\}$. 于是 $F[x]$ 中每个非单位非零(即次数$\geqslant 1$)的首 1 多项式 $f(x)$ 惟一地表成有限个首 1 不可约多项式乘积：

$$f(x)=p_1(x)^{l_1}\cdots p_g(x)^{e_g},$$

其中 $p_i(x)(1\leqslant i\leqslant g)$ 是 $F[x]$ 中不同的首 1 不可约多项式.

设 K/k 为有限扩张($k=F(x)$)，即 K 是函数域. 可以定义 K 中的整元素 α，即 α 是某个首 1 多项式 $f(Y)=Y^n+\alpha_{n-1}Y^{n-1}+\cdots+\alpha_0$ 的根，其中 $\alpha_i\in F[x]$，K 中的整元全体形成环，叫作 K 的整元环，和数论中一样表示成 O_K，它也是狄德金整环，即非零理想均惟一表成有限个素理想之积. O_K 是秩为 $[K:k]$ 的自由 $F[x]$—模，从而也有整基概念.

Hensel 的赋值理论是受函数在一点处台劳展开式的启发，现在又用于研究函数域. 对 $F(x)$ 的每个首 1 不可约多项式 $p=p(x)$，可以与数域情形一样构作 p-adic 指数赋值，即对 $0\neq\alpha\in F(x)$，定义 $v_p(\alpha)=a$，其中 $p^a\parallel\alpha$，而令 $v_p(0)=+\infty$，然后可类似作 p-adic 赋值 $|\alpha|_p=\gamma^{v_p(\alpha)}(0<\gamma<1)$. 于是 $k=F(x)$ 中有无穷多 p-adic 赋值，不同的 p 给出域 k 完全不同的 p-adic 拓朴. 最后，对 $0\neq\alpha\in F[x]$，记 $v_\infty(\alpha)=-\deg\alpha$（多项式 α 的次数），而 $v_\infty\left(\frac{\alpha}{\beta}\right)=v_\infty(\alpha)-v_\infty(\beta)$. 可以证明 v_∞ 满足指数赋值的三个条件，而 $|\alpha|_\infty=\gamma^{v_\infty(\alpha)}(0<\gamma<1)$ 也是 k 上的赋值，叫 ∞-adic 赋值. 注意 k 中超越元 x 的选取方式可以有许多种，如果取 $t=\frac{1}{x}$ 为新的超越元，$k=F(t)$. 对于 $F[x]$ 中多项式可代入 $t=\frac{1}{x}$：

$$
\begin{aligned}
f(x)&=a_0+a_1x+\cdots+a_nx^n \qquad (a_i\in F, a_n\neq 0)\\
&=a_0+\frac{a_1}{t}+\cdots+\frac{a_n}{t^n}=\frac{1}{t^n}(a_0t^n+a_1t^{n-1}+\cdots+a_n).
\end{aligned}
$$

于是 $v_\infty(f)=-n=v_t(f)$，即 ∞-adic 赋值就是 t-adic 赋值. 所以可写 $\infty=\left(\frac{1}{x}\right)$.

对于一般的函数域 K，可象数域情形那样，对 O_K 的每个非零素理想 P 构作 P-adic 赋值（由于 O_K 为狄德金整环），将 K 的子域 k 中超越元 x 换成 $t=\frac{1}{x}$，K 中又可构作出一些新的非阿赋值.

现在我们取 F 为有限域 $\boldsymbol{F}_q(q=p^n, p$ 为素数)，则函数域扩张 $K/k(k=\boldsymbol{F}_q(x))$ 与数域扩张的类比还可以走得更远. 设 P 和 $\mathscr{P}$ 分别为 O_K 和 $O_k=\boldsymbol{F}_q[x]$ 的非零素理想，并且 $P|\mathscr{P}$（即 P 是 O_K 中理想 $\mathscr{P}O_K$ 的素因子，则 $\frac{O_K}{P}$ 是有限域，这是因为 $\frac{O_k}{\mathscr{P}}=\frac{\boldsymbol{F}_q[x]}{(p(x))}$ 是有限域 $\boldsymbol{F}_{q^d}$（$\mathscr{P}$ 为主理想 $(p(x))$，而 $d=\deg p(x)$），并且 $f=\left[\frac{O_K}{P}:\frac{O_k}{\mathscr{P}}\right]$ 有限，f 叫 P 对 $\mathscr{P}$ 的**剩余类域次数**，表示成 $f(P/\mathscr{P})$.

同样地，若 $\mathscr{P}O_K = \prod_P P^{e_p}$，则 e_p 叫 P 对 $\mathscr{P}$ 的分歧指数. 若 $\mathscr{P}O_K = P_1^{e_1}\cdots P_g^{e_g}$，其中 $P_i(1\leqslant i\leqslant g)$ 为 O_K 的不同素理想，我们也有 $\sum_{i=1}^{g} e(P_i/\mathscr{P})f(P_i/\mathscr{P})=[K:k]$. 另一方面，可以证明所有 p-adic 赋值（$p$ 过 $\boldsymbol{F}_p[x]$ 的全部首 1 不可约多项式）和 $\left(\frac{1}{x}\right)$-adic 赋值是 $k=\boldsymbol{F}_q(x)$ 的全部赋值，对应素除子表示成 p 和 ∞. 而 K 的所有赋值为 P-adic 赋值（P 过 O_K 的所有非零素理想）以及由 k 的 $\left(\frac{1}{x}\right)$-adic 赋值扩充到 K 的一些新的赋值，它们都是非阿赋值（即 $\boldsymbol{F}_q$ 为常数域的函数域 K 中没有阿基米德赋值）.

最后，O_K 的分式理想群与主分式理想子群的商叫作 O_K 的理想类群，表示成 $C(O_K)$，可以证明这是有限阿贝尔群，其阶 $h(O_K)$ 叫 K 的理想类数，它的意义和数论中一样：$h(O_K)=1$ 当且仅当 O_K 为惟一因子分解整环(UFD).

在函数域中还有另一个“类数”概念. 让我们以 $k=\boldsymbol{F}_q(x)$ 为例说明这个概念. k 中每个非零有理函数均可写成

$$f=\alpha\cdot\frac{g_1(x)^{e_1}\cdots g_s(x)^{e_s}}{h_1(x)^{\lambda_1}\cdots h_t(x)^{\lambda_t}},$$

其中 $g_i(x)(1\leqslant i\leqslant s)$ 和 $h_j(x)(1\leqslant j\leqslant t)$ 是 $\boldsymbol{F}_q[x]$ 中彼此不同的 $s+t$ 个首 1 不可约多项式. $e_i,\lambda_j\geqslant 1$. 这时，我们称 f 在位 $g_i(x)$ 处有 e_i 阶**零点**，而在位 $h_j(x)$ 处有 λ_j 阶**极点**. 也就是说，对 $\boldsymbol{F}_q[x]$ 的每个首 1 不可约多项式 $p=p(x)$，若 $v_p(\alpha)=e>0$，则称 α 在 p 处有 e 阶零点. 若 $v_p(\alpha)=-e, e>0$，则称 α 在 p 处有 e 阶极点. $\deg p(x)$ 叫作素除子 p 的**次数**，由于 $\frac{\boldsymbol{F}_q[x]}{(p(x))}=\boldsymbol{F}_{q^{\deg p}}$，从而 $\deg p(x)=\left[\frac{\boldsymbol{F}_q[x]}{(p(x))}:\boldsymbol{F}_q\right]$. 最后对 $t=\frac{1}{x}$，它是 $\boldsymbol{F}_q[t]$ 中的一次多项式，即素除子 ∞ 的次数为 1. 而

$$v_t(f)=-\deg(g_1^{e_1}\cdots g_s^{e_s})+\deg(h_1^{\lambda_1}\cdots h_t^{\lambda_t})$$

$$= -\sum_{j=1}^{t} v_{h_j}(f) \cdot \deg h_j - \sum_{i=1}^{s} v_{g_i}(f) \cdot \deg g_i.$$

这就表明 $\prod_P v_p(f)\deg p = 0$，其中 P 过 $k = \boldsymbol{F}_q(x)$ 的所有素除子（包括 $p = \infty$）.

现在对常数域为 $\boldsymbol{F}_q$ 的任意函数域 K，由 K 的所有素除子为基生成的自由阿贝尔群叫作 K 的**除子群**，表示成 $D(K)$. 其中元素唯一表示成

$$A = \prod_P P^{e_p}(\text{有限乘积}),\ e_p \in \mathbf{Z}, \qquad (*)$$

叫作 K 的一个**除子**，当 P 为有限素除子（即 P 是 O_K 的非零素理想）时，定义 P 的次数为 $\deg P = \left[\frac{O_K}{P} : \boldsymbol{F}_q\right]$. 而 P 的范为 $N(P) = \left|\frac{O_K}{P}\right| = |\boldsymbol{F}_q|^{\deg P} = q^{\deg P}$. 对无限素除子可类似定义 $\deg P$ 和 $N(P)$. 对于由（$*$）式表示的除子 A 则定义

$$\deg A = \sum_P e_P \cdot \deg P, N(A) = \prod_P N(P)^{e_P}.$$

则零次除子全体为 $D(K)$ 的子群，记为 $D^0(K)$. 进而，对每个 $0 \neq f \in K$，可以构作一个除子，

$$\operatorname{div}(f) = \prod_P P^{v_P(f)}(\text{有限乘积})$$

称为**主除子**. 可以证明：主除子的次数为零，并且全体主除子构成 $D^0(K)$ 的一个子群，表示成 $P(K)$. 最后，可以证明 $\frac{D^0(K)}{P(K)}$ 是有限阿贝尔群，叫作 K 的**除子类群**，表示成 $C(K)$. 它的阶 $h(K) = |C(K)|$ 叫作 K 的**除子类数**. 于是：$h(K) = 1$ 当且仅当 K 的所有零次除子都是主除子.

于是，在函数域中有两个类群和类数概念，理想类群 $C(O_K)$ 和类数 $h(O_K)$ 来源于数论，用来刻划整数环 O_K 的性质. 而除子类群 $C(K)$ 和类数 $h(K)$ 则来源于函数论和几何学：给了一些正整数

$a_1,\cdots,a_r,b_1,\cdots,b_s$ 和 K 的一些素除子 $P_1,\cdots,P_r,Q_1,\cdots,Q_s$. 使得 $\sum_{i=1}^{r}a_i\deg P_i=\sum_{j=1}^{s}b_j\deg Q_j$. 能否在 K 中找到函数 f,使得 f 在 P_i 处有 a_i 阶零点($1\leqslant i\leqslant r$),在 Q_j 处有 b_j 阶极点($1\leqslant j\leqslant s$),而在其他素除子处既无零点也无极点? 则:答案是肯定的当且仅当 $h(K)=1$.

由数论和函数论引出的两个不同的类数有简单的联系:$h(K)=h(O_K)\cdot R(K)$,其中 $R(K)$叫作函数域 K 的 regulator. 它也是由 O_K 的单位群中基本单位系所决定的,但它是一个整数(而数域的 regulator 通常是一个复杂的无理数). 因此 $h(O_K)|h(K)$. 今后我们主要谈除子类数 $h(K)$.

能否象数域情形找到 $h(K)$的解析计算公式? 数域 K 的类数解析公式是由狄德金 zeta 函数

$$\zeta_K(s)=\sum_{A}N(A)^{-s}$$

在 $s=1$ 处的留数给出的(其中 A 过 O_K 的所有非零整理想). 类似地我们可以构作函数域 K 的 zeta 函数

$$\zeta_K(s)=\sum_{A}N(A)^{-s},$$

其中 A 过 K 的所有整除子,即 $A=P_1^{e_1}\cdots P_g^{e_g}$, e_i 均$\geqslant 1$,由于 $N(A)=q^{\deg A}$,因此

$$\zeta_K(s)=\sum_{A}q^{-s\deg A}=\sum_{n\geqslant 0}a_nq^{-ns}\qquad(\mathrm{Re}(s)>1),$$

其中 a_n 为 K 中次数为 n 的整除子个数. 可以证明这个级数在 $\mathrm{Re}(s)>1$ 时收敛. 由于 $N(AB)=N(A)N(B)$,所以它也可展开成欧拉无穷乘积的形式.

$$\zeta_K(s)=\prod_{P}(1-N(P)^{-s})^{-1}\qquad\mathrm{Re}(s)>1,$$

其中 P 过 K 的所有(非零)素除子,$\zeta_K(s)$是 q^{-s}的函数,令 $U=q^{-s}$,则这个 zeta 函数变成了

$$Z_K(U) = \sum_A U^{\deg A} = \sum_{n\geqslant 0} a_n U^n$$
$$= \prod_P (1 - U^{\deg p})^{-1} \quad (|U| < \frac{1}{q}).$$

现在我们计算 $k=\boldsymbol{F}_q(x)$ 的 zeta 函数. 首先用乘积表达式把无限素除子 $\infty=\left(\frac{1}{t}\right)$ 和有限素除子 $(p(x))$ 分开, 由于 $\deg(\infty)=1$, 所以

$$Z_k(U) = (1-U)^{-1} \cdot \prod_p{}'(1-U^{\deg p})^{-1}$$
$$= \frac{1}{(1-U)} \cdot \sum_A{}' U^{\deg A},$$

其中 Π' 是对有限素除子求积, Σ' 是对 $O_k = \boldsymbol{F}_q[x]$ 的理想 A 求和, 每个理想 A 由首 1 多项式生成. 而 $\boldsymbol{F}_q[x]$ 中 n 次首 1 多项式共 q^n 个, 于是

$$Z_k(U) = \frac{1}{1-U} \cdot \sum_{n\geqslant 0} U^n \cdot q^n = \frac{1}{(1-U)(1-qU)},$$

即 $Z_k(U)$ 是非常简单的有理函数. 而

$$\zeta_k(s) = \frac{1}{(1-q^{-s})(1-q^{1-s})}.$$

现在设 K 是以 $\boldsymbol{F}_q$ 为常数域的任意函数域. 则 $K=\boldsymbol{F}_q(x,y)$, 其中 $f(x,y)=0$, 而 $f(X,Y)\in\boldsymbol{F}_q[X,Y]$. 于是 K 是曲线 $C: f(X,Y)=0$ 在 $\boldsymbol{F}_q$ 上的函数域. 曲线 C 的亏格 $g(C)$ 也叫作函数域 K 的亏格, 表示成 $g(K)$. 利用几何上一个重要结果: 黎曼-洛赫定理, 证明了对任意这种函数域 K, $Z_K(U)$ 均是 U 的有理函数, 并且分母均为 $(1-U)(1-qU)$, 即

$$Z_K(U) = \frac{L(U)}{(1-U)(1-qU)},$$

其中 $L(U)=1+c_1U+c_2U^2+\cdots+c_{2g}U^{2g}\in\boldsymbol{Z}[U]$, 即 c_i 均为有理整数, 并且 $c_{2g}=q^g$, 而 g 是域 K 的亏格 $g(K)$. 由于 $k=\boldsymbol{F}_q(x)$ 是直线的函数域, 而直线的亏格为 0, 即 $g(k)=0$, 因此 $Z_K(U)=$

$\frac{1}{(1-U)(1-qU)}$，与前面计算结果符合.令

$$L(U)=\prod_{i=1}^{2g}(1-w_iU).$$

由于 $L(U)$ 的常数项为 1，而 $w_i(1\leqslant i\leqslant 2g)$ 是多项式 $2g$ 个根的倒数.可知 $w_i(1\leqslant i\leqslant 2g)$ 均是代数整数.

利用黎曼-洛赫定理还证明了：

定理 $L(1)=h(K)$ 即 K 的除子类数为

$$h(K)=1+c_1+\cdots+c_{2g-1}+q^g=\prod_{i=1}^{2g}(1-w_i).$$

由于 $h(K)$ 是正整数，由 $Z_k(U)=L(U)/(1-U)(1-qU)$ 可知 $U=1$ 是 $Z_K(U)$ 的单极点，并且留数为

$$\operatorname*{res}_{U=1}Z_K(U)=\lim_{U\to 1}(U-1)Z_K(U)=\frac{L(1)}{q-1}=\frac{h(K)}{q-1}.$$

再由 $h(K)=R(K)h(O_K)$，可知

$$R(K)h(O_K)=h(K)=(q-1)\cdot\operatorname*{res}_{U=1}Z_K(U).$$

如果能把 $Z_K(U)$ 在 $U=1$ 的留数算出，便给出 $h(K)$（或 $R(K)h(O_K)$）的解析公式.与数域情形类似，理想类数 $h(O_K)$ 和 regulator $R(K)$ 仍旧混在一起. 1924 年，阿廷的博士论文就是用此法得到二次函数域 K（即 $[K:k]=2$）的类数公式.

以上叙述的函数域结果是在 20 世纪前 30 年作出的，这些结果与数域情形非常相似，即是函数域的算术理论.从此之后，有限域上单变量代数函数域（今后简称函数域，几何上它是有限域上代数曲线的函数域）和数域常常同时（或平行地）加以研究，它们统称为整体域，它们对某个素除子的局部化都叫作局部域.比如说，关于二次型的哈瑟定理对于函数域也成立.即：若 $f(x_1,\cdots,x_k)=\sum_{i,j=1}^{k}a_{ij}x_ix_j$ 是函数域 K 上的二次型（$a_{ij}\in K$），则对每个 $0=\alpha\in$

K,方程 $f(x_1,\cdots,x_k)=\alpha$ 在 K 中有解当且仅当对 K 的每个素除子 P,方程在局部域 K_P 中均有解.类似地,f 在 K 中可表示 0(即 $f(x_1,\cdots,x_k)=0$ 在 K 中有非零解)当且仅当 f 在每个局部域 K_P 中可表示 0.

函数域不仅是数论对象,也是几何对象(它是代数曲线的函数域),所以希望对函数域算术性质的研究能得到代数几何学的结果.人们经过 20 世纪前 30 年的实例考查,终于在 40 年代得到重大成果.而这个成果和希尔伯特的一个问题有关.

参 考 文 献

[1] E. Artin, Algebraic Numbers and Algebraic Functions, 1951.

[2] H. Hasse, Number Theory, (第三版) Akademie-Verlag, 1979.(详细讲述数域和函数域的数论性质.)

§3.4 韦依定理——函数域上的黎曼猜想(20 世纪 40 年代)

希尔伯特第 8 问题是:在有理数域和有理整数环 **Z** 上的数论问题(如黎曼猜想,素数分布和哥德巴赫问题等)能否推广到任意数域 K 和它的整数环 O_K 上?

20 世纪初期,德国数学家 Landau 等人使用解析数论工具,得

到$\zeta_K(s)$的一些与黎曼 zeta 函数$\zeta(s)(=\zeta_Q(s))$类似的解析性质，把素数分布的重要结果平行地推广成O_K中素理想分布结果. 对于每个代数数域K，人们也提出了相应的(广义)黎曼猜想，即$\zeta_K(s)$的非平凡零点都在直线$\mathrm{Re}(s)=\frac{1}{2}$之上. 但是目前对任何一个数域(包括$Q$)都没有解决. 可是在20世纪40年代，函数域上类似的黎曼猜想却得到证明.

由于$\zeta_K(s)$(K为函数域)是q^{-s}的有理函数，所以解析延拓已不成问题. 利用黎曼-洛赫定理可以得到$\zeta_K(s)$的函数方程：

$$\zeta_K(s)=q^{(g-1)(1-2s)}\zeta_K(1-s).$$

这也相当于

$$Z_K(U)=(\sqrt{q}\,U)^{2g-2}Z_K\left(\frac{1}{qU}\right).$$

由于

$$Z_K(U)=\frac{L(U)}{(1-U)(1-qU)},$$
$$L(U)=1+c_1U+\cdots+c_{2g}U^{2g}$$
$$=\prod_{i=1}^{2g}(1-w_iU)\in \mathbf{Z}[U],$$

从而得到$2g$次多项式$L(U)$的函数方程

$$L(U)=(\sqrt{q}\,U)^{2g}L\left(\frac{1}{qU}\right).$$

于是得到多项式系数(有理整数)c_n和根的倒数w_i有关系

$$c_{2g-i}=q^{g-i}c_i \quad (0\leqslant i\leqslant g) \quad (c_0=1, c_{2g}=q^g).$$

从而$Z_K(U)$和$\zeta_K(s)$由g个有理整系数$c_1,c_2,\cdots,c_g$所完全决定. 而

$$\prod_{i=1}^{2g}(1-w_iU)=\prod_{i=1}^{2g}\left(1-\frac{q}{w_i}U\right).$$

从而适当调整w_i的次序，可以使得

$$w_{2g-i+1}w_i=q \ (1\leqslant i\leqslant g).$$

即$2g$个w_i可分成g对，每对w_i相乘均为q.

接下来便是韦依(A. Weil,1906—1998)的一个重要猜想. 1939 年二战爆发前夕,韦依在法国斯特拉斯堡教书,和 H. Cartan,J. Dieudonné 等人创建了布尔巴基数学小组. 二战爆发后遭受许多惊险(被指控为苏联间谍关进监狱并差一点被枪毙),于 1940 年到了美国,在美国研读高斯的两篇文章,研究有限域上费马曲线 $x^n+y^n=1$ 和 Artin-Schreier 曲线 $y^q+y=f(x)(f(x)\in \boldsymbol{F}_q[x])$的函数域. 用高斯和计算出这些函数域的 zeta 函数. 在这些计算的基础上他于 1940 年的一篇短文中提出如下的猜想:

韦依猜想 对常数域为 $\boldsymbol{F}_q$ 的每个函数域 K,均有 $|w_i|=\sqrt{q}$ $(1\leqslant i\leqslant 2g)$,其中 $g=g(K)$为 K 的亏格.

由于 $w_i^{-1}(1\leqslant i\leqslant 2g)$是多项式 $L(U)$的全部零点,从而也是 $Z_K(U)=L(U)/(1-U)(1-qU)$的全部零点. 但是 $U=q^{-s}$. 令 $q^{-s}=w_i^{-1}$,$s=\sigma+it(\sigma=\mathrm{Re}(s))$,则 $|w_i|=|q^s|=q^\sigma$. 从而韦依猜想相当于 $q^\sigma=\sqrt{q}$,即 $\mathrm{Re}(s)=\frac{1}{2}$. 所以韦依猜想是说:函数域 K 的 zeta 函数 $\zeta_K(s)$所有($2g$ 个)零点都在直线 $\mathrm{Re}(s)=\frac{1}{2}$之上. 这正是黎曼猜想在函数域上的模拟!

经过几年的奔波(其间还在巴西工作了两年),韦依于 1947 年才找到与他水平相称的职位:芝加哥大学数学教授. 1948 年韦依证明了他的上述猜想. 在介绍他的证明之前,我们先说明韦依猜想的几何意义. 事实上,韦依在计算上述曲线的 zeta 函数 $Z_K(U)$时,并没有采用我们在上节介绍的数论定义 $Z_K(U)=\sum\limits_{A}U^{\deg A}$,他发现 $Z_K(U)$还有一个几何定义,随后用代数几何方法证明的猜想.

有限域 $\boldsymbol{F}_q$ 为常数域的每个函数域 K 都是 $\boldsymbol{F}_q$ 上某条曲线 C: $f(X,Y)=0$ 的函数域:$K=\boldsymbol{F}_q(C)=\boldsymbol{F}_q(x,y)$,其中 $f(x,y)=0$(而 $f(X,Y)\in \boldsymbol{F}_q[X,Y]$). 彼此双有理等价的曲线给出同一个函数域

(或者更正确地说:给出彼此同构的函数域). 所以我们不妨设曲线 C 是充分"好"的,即是一条"绝对不可约的无奇点"曲线(每条曲线均双有理等价于这种"好"曲线,这是代数几何中的一个很不简单的结果).

于是我们转向几何,考虑曲线 C 在 $\boldsymbol{F}_q^2$ 中的点数,即方程 $f(X,Y)=0$在 $\boldsymbol{F}_q$ 中的解数(有限域中的方程只有有限多解). 但是还要向前走一步,我们不是考虑方程在仿射平面 $\boldsymbol{F}_q^2$ 中的解,而是在射影平面 $\boldsymbol{P}^2(\boldsymbol{F}_q)$中的解. 射影平面比仿射平面增加一些"无穷远点",确切说来,对于彼此平行的仿射直线都增加一个公共的无穷远点而变成射影直线. 由于仿射平面 $\boldsymbol{F}_q^2$ 中的直线共有 $q+1$ 个不同的方向(斜率),所以共增加 $q+1$ 个无穷远点,这 $q+1$ 个无穷远点又组成一条射影直线,所以射影平面 $\boldsymbol{P}^2(\boldsymbol{F}_q)$共有 q^2+q+1 个点,每条射影直线上都有 $q+1$ 个点.

仿射几何的许多论断用射影几何语言则更加漂亮. 在仿射平面中,两条不同仿射直线有相交和平行两种可能性,但是在射影平面中,任意两条不同的射影直线都恰好相交于一点. 对我们现在的问题,也是考虑曲线 C 在 $\boldsymbol{P}^2(\boldsymbol{F}_q)$中的射影点数,即把 C 上的无穷远点也考虑在内. 对每个正整数 n,我们以 $N_n(C)$表示曲线 C 在 $\boldsymbol{P}^2(\boldsymbol{F}_{q^n})$中的射影点数. 定义

$$Z_c(U)=\exp\left(\sum_{n\geqslant 1}\frac{N_n(C)}{n}U^n\right),$$

其中 $\exp(x)=e^x$ 为指数函数,$Z_c(U)$叫作**曲线** C **的 zeta 函数**.

以直线 L 为例,射影直线在 $\boldsymbol{F}_{q^n}$上共有 q^n 个点,即 $N_n(L)=q^n+1$. 于是

$$\begin{aligned}Z_L(U)&=\exp\left(\sum_{n\geqslant 1}\frac{q^n+1}{n}U^n\right)\\&=\exp\left(\sum_{n\geqslant 1}\frac{(qU)^n}{n}+\sum_{n\geqslant 1}\frac{U^n}{n}\right)\\&=\exp(-\log(1-qU)-\log(1-U))\end{aligned}$$

$$= \frac{1}{(1-U)(1-qU)}.$$

而这正是直线的函数域 $k=\boldsymbol{F}_q(x)$ 的 zeta 函数 $Z_k(U)$.

利用代数几何方法(本质上系用赋值论,不过几何上把每个赋值等价类称之为"位"),可以证明:若 C 是 $\boldsymbol{F}_q$ 上一条"好"的代数曲线,K 是 C 的函数域,则

$$Z_c(U)=Z_K(U).$$

数论与几何在这里又汇合在一起!韦依计算了 $\boldsymbol{F}_q$ 上费马曲线和 Artin-Schreier 曲线在所有扩域 $\boldsymbol{F}_{q^n}$ 上的射影点数,结果发现 $Z_c(U)$ 的零点都是高斯和(和雅柯比和),它们是分圆域中的代数整数,熟知高斯和的绝对值均为$\sqrt{q}$,由此对这些曲线(以及它们的函数域)验证了猜想的正确性. 1924年,Artin 对二次函数域就提出了这个猜想. 当$2\nmid q$ 时,二次函数域均有形式 $K=k(\sqrt{f(x)})$,其中 $k=\boldsymbol{F}_q(x)$,即它是曲线 $Y^2=f(x)$ 的函数域,$f(x)\in\boldsymbol{F}_q[x]$. 令 $d=\deg f(x)$. 则此曲线(和二次函数域 K)的亏格为$\left[\frac{d-1}{2}\right]$. 1935年,哈瑟和 Davenport 对于亏格为 1(即 $d=3$ 和 4)的曲线(叫作椭圆曲线)证明了这个猜想.

韦依在 1940 年的一篇文章中叙述了他对证明这个猜想的想法,认为这个代数曲线的结果需要用代数曲面上的工具,并且要把意大利几何学派的成果加以改造和精确化. 为了证明这个猜想,他在 1946 年专门写了一本书:《代数几何基础》. 正如 J. -P. Serre 所说(1999 年:《韦依的生平和著作》):"这本三百页的巨著很难懂,而在 20 年后,它又被 Grothendieck 更巨大更难懂的《代数几何原理》所替代. "韦依在书中充分使用了 E. Noether 等人发展的交换代数理论和语言,引入了代数几何一批重要的概念(Cycle,一般点和特殊化,相交重数,曲面上的对应,…),是代数几何发展的一个里程碑,把代数几何研究推进到一个新时期.

在这本书出版之后,韦依证明了他的猜想,成为有限域上代数

曲线的韦依定理. 后来,许多数学家致力于寻求简单的证法,力图避免过多的代数几何知识. 1979 年,美国数学会组织了一次关于希尔伯特问题的专门研讨会,由各方专家报告 23 个希尔伯特问题的进展情况. Bombieri 在文集中刊登了他的报告,给出了韦依定理一个最简单的证明,只用到黎曼-洛赫定理.

韦依定理有许多有趣的应用. 将 $Z_K(U)=Z_c(U)$的两边定义写出来,便得到

$$\frac{\prod_{i=1}^{2g}(1-w_iU)}{(1-U)(1-qU)}=\exp\left(\sum_{n\geq 1}\frac{N_n(C)}{n}U^n\right).$$

两边取对数便得到

$$\sum_{n\geq 1}\frac{N_n(C)}{n}U^n=\log\frac{\prod_{i=1}^{2g}(1-w_iU)}{(1-U)(1-qU)}$$
$$=\sum_{n\geq 1}\frac{1}{n}\left(1+q^n-\sum_{i=1}^{2g}w_i^n\right)U^n.$$

于是对每个 $n\geq 1$,曲线 C 在 $\boldsymbol{F}_{q^n}$中的射影点数有公式

$$N_n(C)=1+q^n-\sum_{i=1}^{2g}w_i^n.$$

由函数方程知 $2g$ 个 w_i 分成 g 对,每对相乘均为 q. 由韦依定理又知 $|w_i|=\sqrt{q}\ (1\leqslant i\leqslant 2g)$,从而其中有 g 个 w_i 为$\sqrt{q}\,e^{i\theta_\lambda}(1\leqslant\lambda\leqslant g)$,而另 g 个为$\sqrt{q}\,e^{-i\theta_\lambda}(1\leqslant\lambda\leqslant g)$. 于是

$$N_n(C)=1+q^n-2\sqrt{q^n}\sum_{i=1}^{g}\cos n\theta_\lambda\quad(n=1,2,3,\cdots).$$

我们只要算出 $\boldsymbol{F}_q$ 的 g 个不同有限扩域中曲线 C 的射影点数,由上边形式的 g 个方程就可算出 $\theta_\lambda(1\leqslant\lambda\leqslant g)$的值. 从而给出在任意有限域 $\boldsymbol{F}_{q^n}$中曲线 C 射影点数 $N_n(C)$的计算公式. 进而,

$$|N_n(C)-(1+q^n)|=2\sqrt{q^n}\sum_{i=1}^{g}\cos n\theta_\lambda$$

$$\leqslant 2g\sqrt{q^n} \quad (n \geqslant 1).$$

当 q 很大时，$2g\sqrt{q^n}$ 比 $1+q^n$ 的阶要小. 这个不等式表明：当 q 很大时，对于任何“好”的代数曲线 C，它在有限域 $\boldsymbol{F}_q$ 中的射影点数都与最简单的曲线——直线上的射影点数 $1+q$ 相差无几（相差不超过 $2g\sqrt{q}$）. 而且还可以从这一连串不等式 $|N_n(C)-(1+q^n)| \leqslant 2g\sqrt{q^n}$ $(n=1,2,3,\cdots)$ 也可推出韦依定理. 所以也可把这串不等式叫作韦依定理，它在数论中有许多重要的应用（三角和估计…），并且在信息科学中也有许多实际应用（将在第五章介绍）.

1948 年，韦依对某些高维代数簇（Grasmann 代数簇等）在有限域上的点数作了具体计算，并在拓朴学的启发下提出了高维代数簇的 zeta 函数零点猜想. 这个猜想一直到 1974 年才被 P. Deligne 证明，使用了更现代的代数几何（就是 Serre 称之为更难懂的 Grothendieck 代数几何），我们于第四章再作简略介绍.

参考文献

（除以下二篇之外，还见 § 4.4 节参考文献）

[1] A. Weil, On the Riemann hypothesis in function fields, Proc. Nat. Acad. Sci. USA 27(1941), 345～347.

[2] A. Weil, Numbers of solutions of equations in finite fields, Bull. Amer. Math. Soc. 55(1949), 497～508.



图 12　爱森斯坦

§3.5　模形式理论

1847 年,爱森斯坦(Eisenstein,1823—1852)在 120 页的长文《椭圆函数的报告》中研究椭圆函数的解析性质和数论应用. 而克

罗内克(Kronecker,1823—1891)从1853年起对椭圆函数作了40年的研究,也带有明确的数论目的,就是想完成他的青春之梦:椭圆函数的特殊值添加到虚二次数域 K 上,可构作出 K 的最大阿贝尔扩域.

椭圆函数 $f(z)$ 是复平面上的双周期亚纯函数,即存在两个非零复数 $w_1, w_2, \frac{w_2}{w_1} \notin \boldsymbol{R}$(即 o, w_1, w_2 不在一条直线上),使得 $f(z) = f(z+w_1) = f(z+w_2)$(对每个 $z \in \boldsymbol{C}$). 令

$$L = \boldsymbol{Z}w_1 \oplus \boldsymbol{Z}w_2 = \{n_1 w_1 + n_2 w_2 \mid n_1, n_2 \in \boldsymbol{Z}\},$$

则对每个 $w \in L, z \in \boldsymbol{C}$,均有 $f(z+w) = f(z)$. 于是 $f(z)$ 可看成是 $\boldsymbol{C}/L$ 上的(亚纯)复变函数. L 是秩2的(加法)自由 $\boldsymbol{Z}$-模,叫作 $\boldsymbol{C}$ 的一个格. 由于 $\boldsymbol{C}/L$ 是紧的黎曼面,使得椭圆函数有许多特别的性质.

除了常值函数之外,第一个非平凡的椭圆函数的例子是由外尔斯特拉斯(Weierstrass)给出的:

$$\mathscr{P}(z) = \frac{1}{z^2} + \sum_{0 \neq w \in L} \left(\frac{1}{(z-w)^2} - \frac{1}{w^2} \right),$$

$$\mathscr{P}'(z) = -2 \sum_{w \in L} (z-w)^{-3}.$$

它们在 $z=0$ 处的展开式为

$$\mathscr{P}(z) = \frac{1}{z^2} + \sum_{n=1}^{\infty} (2n+1) s_{2n+2}(L) z^{2n},$$

$$\mathscr{P}'(z) = -\frac{1}{z^3} + \sum_{n=1}^{\infty} 2n(2n+1) s_{2n+2}(L) z^{2n-1},$$

其中

$$s_k(L) = \sum_{0 \neq w \in L} w^{-k} \qquad (k \geqslant 3)$$

是依赖于格 L 的复数(当 $k \geqslant 3$ 时级数收敛,但是 k 为奇数时,易知 $s_k(L) = 0$).

所有以 L 为周期格的椭圆函数的四则运算仍是这样的函数,

从而它们形成域 K_L(关于 L 的椭圆函数域). 可以证明: $K_L=\mathbf{C}(\mathscr{P},\mathscr{P}')$, 即关于 L 的所有椭圆函数都是 $\mathscr{P}$ 和 $\mathscr{P}'$ 在 $\mathbf{C}$ 上的有理函数, 而 $\mathscr{P}'$ 和 $\mathscr{P}$ 是代数相关的: $\mathscr{P}'^2(z)=4\mathscr{P}(z)^3-g_2\mathscr{P}(z)-g_3$, 其中

$$g_2=g_2(L)=60s_4(L),g_3=g_3(L)=140s_6(L).$$

于是关于 L 的椭圆函数域就是 $\mathbf{C}$ 上曲线 $y^2=4x^3-g_2x-g_3$ 的函数域 $\mathbf{C}(x,y)$. 这种曲线是**椭圆曲线**(历史上, 这个名称的来源是: 欧拉和雅柯比在计算椭圆弧长时出现对函数 $y=\sqrt{f(x)}$ 的积分, 其中 $f(x)$ 为三次(和四次)多项式). 所以椭圆函数和即将介绍的模形式, 与下节要讲的椭圆曲线从一开始就有密切的联系.

$\mathscr{P}(L)$ 的系数 $s_k(L)$ 有如下的齐次性质: 若 L 和 L' 是两个**相似**的格, 即有 $\alpha\in\mathbf{C}^\times$ 使得 $L=\alpha L'$, 则

$$s_k(L)=\sum_{0\neq w\in L}w^{-k}=\sum_{0\neq w\in L'}(\alpha w)^{-k}=\alpha^{-k}s_k(L').$$

取 L 的一组 $\mathbf{Z}$-基 w_1,w_2, 即 $L=\mathbf{Z}w_1\oplus\mathbf{Z}w_2$, 则

$$s_k(L)=\sum_{n,m\in\mathbf{Z}}{}'(nw_1+mw_2)^{-k},$$

其中 Σ' 表示求和中去掉 $(n,m)=(0,0)$. 必要时交换 w_1 和 w_2 的次序, 可设 $\frac{w_2}{w_1}=\tau$ 的虚数部分为正, 即 $\tau\in\mathbf{H}$, 其中

$$\mathbf{H}=\{\alpha\in\mathbf{C}\mid I_m(\alpha)>0\}$$

叫作**上半平面**. 于是 $s_k(L)=w_1^{-k}G_k(\tau)$, 其中

$$G_k(\tau)=s_k(\mathbf{Z}\oplus\mathbf{Z}\tau)=\sum_{n,m\in\mathbf{Z}}{}'(n\tau+m)^{-k}\qquad(\tau\in\mathbf{H})$$

叫作**爱森斯坦级数**$(2|k\geqslant4)$. 格 L 的 $\mathbf{Z}$-基有不同选取方式. 易知: w'_1,w'_2 为 L 的另一组 $\mathbf{Z}$-基, 当且仅当存在行列式为 ±1 的整数方阵 $\begin{pmatrix}a&b\\c&d\end{pmatrix}$, 使得

$$\begin{pmatrix}w'_2\\w'_1\end{pmatrix}=\begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}w_2\\w_1\end{pmatrix}=\begin{pmatrix}aw_2+bw_1\\cw_2+dw_1\end{pmatrix}.$$

于是令 $\tau=\frac{w_2}{w_1}, \tau'=\frac{w'_2}{w'_1}$，则 $\tau'=\frac{a\tau+b}{c\tau+d}$. 如果 $\tau\in\boldsymbol{H}$，则 $\tau'\in\boldsymbol{H}$ 当且仅当 $ad-bc=1$，并且这时

$$G_k(\tau')=(w'_1)^k s_k(L)=\left(\frac{w'_1}{w_1}\right)^k G_k(\tau)=(c\tau+d)^k G(\tau).$$

于是我们有 $\boldsymbol{H}$ 上的解析函数 $G_k(\tau)$. 模群

$$\Gamma=SL_2(\mathbf{Z})=\left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a,b,c,d\in\mathbf{Z}, ad-bc=1\right\}$$

按如下方式作用在 $\boldsymbol{H}$ 上：对 $r=\begin{pmatrix} a & b \\ c & d \end{pmatrix}\in\Gamma$，

$$\gamma(\tau)=\frac{a\tau+b}{c\tau+d} \qquad (\tau\in\boldsymbol{H}).$$

而 $G_k(\tau)$ 在这种作用下有性质

$$G_k(\gamma(\tau))=(c\tau+d)^k G_k(\tau).$$

由于当 k 为奇数时，$G_k(\tau)=0$. 以下考虑 $G_{2k}(\tau)(k\geqslant 2)$. 取 $\begin{pmatrix} 11 \\ 01 \end{pmatrix}\in SL_2(\mathbf{Z})$，则 $G_{2k}(\tau+1)=G_{2k}(\tau)$，即 $G_{2k}(\tau)$ 是周期为 1 的函数. 从而有 Fourier 展开

$$G_{2k}(\tau) = \sum_{m,n\in\mathbf{Z}}{}'(m\tau+n)^{-2k} = \sum_{n\in\mathbf{Z}} c_n q^n$$
$$(q = e^{2\pi i\tau}, \tau\in\boldsymbol{H}).$$

经过计算可知上式右边当 $n<0$ 时 $c_n=0$，即

$$G_{2k}(\tau) = 2\zeta(2k) + \frac{2(2\pi i)^{2k}}{(2k-1)!}\sum_{n=1}^{\infty}\sigma_{2k-1}(n)q^n \quad (k\geqslant 2),$$

其中 $\sigma_{2k-1}(n)=\sum_{d\mid n}d^{2k-1}$（$n$ 的所有正因子的 $2k-1$ 次方之和）. 由于 $\zeta(2k)=-\frac{(2\pi i)^{2k}}{2(2k)!}B_{2k}$，可知

$$G_{2k}(\tau)=2\zeta(2k)E_{2k}(\tau),$$

其中 $E_{2k}(\tau)=1-\frac{4k}{B_{2k}}\sum_{n=1}^{\infty}\sigma_{2k-1}(n)q^n$ 的所有系数都是有理数. $G_{2k}(\tau)$ 中 c_n 对负整数 n 均为 0 意味着此函数在 $q=0$ 处（即 $\tau=$

$+i\infty$处)解析. 具有上述性质的函数叫作模形式. 确切地说,

定义 H 上复解析函数 $f(\tau)$ 叫作对于群 $\Gamma=SL_2(\mathbf{Z})$ 的权 k 模形式($k\geqslant 0$),是指

(1) 对每个 $\gamma=\begin{pmatrix} a & b \\ c & d \end{pmatrix}\in\Gamma$, $f\left(\dfrac{a\tau+b}{c\tau+d}\right)=(c\tau+d)^k f(\tau)$

$(\tau\in \boldsymbol{H})$;

(2) $f(\tau)$ 在 $\tau=+i\infty$ 处解析,即有 Fourier 展开

$$f(\tau)=\sum_{n=0}^{\infty}c_n q^n \quad (q=e^{2\pi i\tau},\tau\in \boldsymbol{H}).$$

当 $c_0=0$ 时, $f(\tau)$ 叫作尖点(cusp)模形式. 以 $M_k(\Gamma)$ 和 $S_k(\Gamma)$ 分别表示对 Γ 的全体权 k 模形式和尖点模形式所成的集合,这是 $\boldsymbol{C}$ 上的向量空间. 可以证明它们均是有限维向量空间,并且维数是

$$\dim(M_{2k}(\Gamma))=\begin{cases}\left[\dfrac{k}{6}\right], 若\ k\equiv 1 \pmod 6,\\ \left[\dfrac{k}{6}\right]+1, 若\ k\not\equiv 1 \pmod 6.\end{cases}$$

(取 $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\in\Gamma$, 则 $f(\tau)=(-1)^k f(\tau)$, 可知当 k 为奇数时, $M_k(\Gamma)$ 中只有零函数)由定义易知,若 $f\in M_k(\Gamma)$, $g\in M_{k'}(\Gamma)$, 则 $f_g\in M_{k+k'}(\Gamma)$.

例 易知 $M_0(F)=\boldsymbol{C}$, 由维数公式知 $M_2(\Gamma)=\{0\}$. 由于维数公式和爱森斯坦级数

$$E_4(\tau)=1+240\sum_{n=1}^{\infty}\sigma_3(n)q^n\in M_4(\Gamma),$$

$$E_6(\tau)=1-504\sum_{n=1}^{\infty}\sigma_5(n)q^n\in M_6(\Gamma),$$

可知 $M_4(\Gamma)=\boldsymbol{C}E_4(\tau)$, $M_6(\Gamma)=\boldsymbol{C}E_6(\tau)$, 于是 $M_8(\Gamma)=\boldsymbol{C}E_4^2(\tau)=\boldsymbol{C}E_8(\tau)$, $M_{10}(\Gamma)=\boldsymbol{C}E_4(\tau)E_6(\tau)=\boldsymbol{C}E_{10}(\tau)$, 并且 $\Delta(\tau)=E_4(\tau)^3-$

$E_6(\tau)^2 \in s_{12}(\Gamma)$.

可以证明当 $k<6$ 时，$S_{2k}(\Gamma)=\{0\}$，而当 $k\geqslant 6$ 时，$S_{2k}(\Gamma)=\Delta M_{2k-12}(\Gamma)$. 并且 $M_{2k}(\Gamma)=S_{2k}(\Gamma)\oplus \boldsymbol{C}E_{2k}(\tau)$.

类似地，可以对 $\Gamma=SL_2(\boldsymbol{Z})$ 的某些子群 G，定义 $\boldsymbol{H}$ 上关于群 G 的权 k 模形式和尖点模形式，它们构成 $\boldsymbol{C}$ 上向量空间分别表示成 $M_k(G)$ 和 $S_k(G)$. 在数论上应用最多的子群 G 为（对正整数 N）

$$\Gamma(N)=\left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix}\in\Gamma \,\middle|\, \begin{pmatrix} a & b \\ c & d \end{pmatrix}\equiv\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}(\bmod N)\right\},$$

$$\Gamma_1(N)=\left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix}\in\Gamma \,\middle|\, \begin{pmatrix} a & b \\ c & d \end{pmatrix}\equiv\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}(\bmod N)\right\},$$

$$\Gamma_0(N)=\left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix}\in\Gamma \,\middle|\, c\equiv 0(\bmod N)\right\}.$$

当 $N\geqslant 2$ 时，$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ 不属于上述子群. 所以当 k 为奇数时，关于上述子群的权 k 模形式可能有非零函数. 但是对于上述子群 G，$M_k(G)$ 均是有限维向量空间，我们已知这些空间的维数. 模形式理论的基本问题是：如何构作向量空间 $M_k(G)$ 的一组基？这个问题至今没有完全解决（特别是奇权情形），对某些模形式空间人们还没有找到足够多的模形式.

除了对椭圆函数的研究之外，模形式理论的另一个数论背景为二次型的整数解问题，即希尔伯特第 11 问题关于整数解的情形. 以有理数域 $\boldsymbol{Q}$ 为例，设

$$f(x_1,\cdots,x_k)=\sum_{i,j=1}^{k}a_{ij}x_ix_j \qquad (a_{ij}\in\boldsymbol{Z})$$

是有理整系数的正定二次型，对每个正整数 n，方程 $f(x_1,\cdots,x_k)=n$ 只能有有限多组有理整数解，以 $N_f(n)$ 表示这个解数（$N_f(n)=0$ 时该方程无整数解），我们能否给出 $N_f(n)$ 的一般表达式（它是 n 的函数）？为此，考虑 $N_f(n)$ 的母函数

$$\theta_f(\tau)=\sum_{x_1,\cdots,x_k\in\boldsymbol{Z}}e^{f(x_1,\cdots,x_k)2\pi i\tau}=\sum_{n=0}^{\infty}N_f(n)e^{2\pi in\tau}\quad(\tau\in\boldsymbol{H})$$

对于许多二次型 f，$\theta_f(\tau)$是对某个子群 G 的权$\frac{k}{2}$模形式. 如果我们知道模形式空间 $M_{k/2}(G)$的一组基

$$\varphi_\lambda(\tau) = \sum_{n\geqslant 0}^{\infty} c_{\lambda n} q^n \quad (q = e^{2\pi i\tau}, \tau \in \boldsymbol{H}) \qquad (1 \leqslant \lambda \leqslant m),$$

其中 $m = \dim M_{k/2}(G)$. 即这些模形式的 Fourier 系数 $C_{\lambda n}$是已知的，则 $\theta_f = \alpha_1\varphi_1 + \cdots + \alpha_m\varphi_m \quad (\alpha_i \in \boldsymbol{C})$. 于是

$$N_f(n) = \alpha_1 c_{1n} + \alpha_2 C_{2n} + \cdots + \alpha_m c_{mn} \quad (n = 0, 1, 2, \cdots).$$

对某些(小的)n，具体算出 $f(x_1, \cdots, x_k) = n$ 的整解个数 $N_f(n)$是不困难的. 假设我们算出 $N_f(0), \cdots, N_f(m-1)$(显然当 f 正定时 $N_f(0) = 1$)，则有关于 $\alpha_1, \cdots, \alpha_m$ 的线性方程组 $N_f(n) = \alpha_1 c_{1n} + \cdots + \alpha_m c_{mn} (0 \leqslant n \leqslant m-1)$. 由此解出 $\alpha_1, \cdots, \alpha_m$ 的值，于是对任意正整数 n，便得到一般公式 $N_f(n) = \alpha_1 c_{1n} + \cdots + \alpha_m c_{mn}$.

例如对平方和问题

$$x_1^2 + x_2^2 + \cdots + x_k^2 = n \qquad (n \geqslant 0),$$

以 $N_k(n)$表示此方程的整解个数. 利用黎曼在证明 $\zeta(s)$函数方程中所使用的 theta 函数

$$\theta(\tau) = \sum_{n\in\mathbf{Z}} q^{n^2} = \sum_{n\in\mathbf{Z}} e^{2\pi i n^2\tau} \qquad (\tau \in \boldsymbol{H}),$$

易知

$$\theta^k(\tau) = \sum_{n=0}^{\infty} N_k(n) q^n.$$

可以证明：当 $4 | k \geqslant 4$ 时，$\theta^k(\tau) \in M_{k/2}(\Gamma_0(4))$.

历史上，二平方和解数 $N_2(n)$的一般公式是由高斯在 19 世纪用环 $\mathbf{Z}[i]$的因子分解特性得到的. 用代数方法高斯还得到 $N_3(n)$和 $N_4(n)$的一般公式，其中 $N_3(n)$是采用三元二次型分类和约化方法，而 $N_4(n)$则使用四元数体的代数结构. 另一方面，欧拉和雅柯比(C. G. Jacobi，1804—1851)用解析方法直接算出 $\theta^k(\tau)(k=2, 3, 4)$的 Fourier 系数 $N_k(n)$，得到同样的结果. 当 $k \geqslant 5$ 时没有适宜的代数结构用来计算 $N_k(n)$，而解析方法也变得复杂. 到了 20 世

纪采用模形式理论，对于满足 $4|k$ 的较小 k 值构作出模形式空间 $N_{k/2}(\Gamma_0(4))$ 的一组基，才得到 $N_8(n),N_{12}(n),\cdots$ 的一般公式（随着 k 增大公式也愈加复杂）. 此后对于 $k\equiv 2 \pmod 4$ 情形，克服更多的困难给出奇权 $k/2$ 模形式空间的基，得到 $N_6(n),N_{10}(n),\cdots$ 的一般公式，而对于 k 为奇数的情形，需要研究半整权 $\left(\frac{k}{2}\in\frac{1}{2}\boldsymbol{Z}\right)$ 模形式. 半整权模型式的系统理论是在 20 世纪 70 年代才由志村五郎(Shimura)给出的，所以 $N_5(n),N_7(n),\cdots$ 的一般公式出现得更晚.

举一个难得的例子. 令 $f(x_1,\cdots,x_8)=\frac{1}{2}\sum_{i,j=1}^{8}a_{ij}x_ix_j$，其中

$$
(a_{ij})=\begin{bmatrix}
2 & 0 & -1 & 0 & 0 & 0 & 0 & 0\\
0 & 2 & 0 & -1 & 0 & 0 & 0 & 0\\
-1 & 0 & 2 & -1 & 0 & 0 & 0 & 0\\
0 & -1 & -1 & 2 & -1 & 0 & 0 & 0\\
0 & 0 & 0 & -1 & 2 & -1 & 0 & 0\\
0 & 0 & 0 & 0 & -1 & 2 & -1 & 0\\
0 & 0 & 0 & 0 & 0 & -1 & 2 & -1\\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 2
\end{bmatrix}
$$

则 f 是整系数多项式，可以证明对应的 theta 函数 θ_f 属于 $M_4(\Gamma)=\boldsymbol{C}E_4$. 于是

$$
\sum_{n=0}^{\infty}N_f(n)q^n=\theta_f(\tau)=cE_4(\tau)
$$

$$
=c\left(1+240\sum_{n=1}^{\infty}\sigma_3(n)q^n\right)\ (q=e^{2\pi i\tau}).
$$

其中 $c\in\boldsymbol{C}$. 由于 (a_{ij}) 是正定阵，方程 $f(x_1,\cdots,x_8)=0$ 只有零解，即 $N_f(0)=1$. 比较上式两边常数项系数，可知 $c=1$，于是

$$
N_f(n)=240\sigma_3(n)=240\sum_{d|n}d^3.
$$

更一般地，如果 $f(x_1,\cdots,x_k)=\frac{1}{2}\sum_{i,j=1}^{k}a_{ij}x_ix_j$，其中 k 阶整数方阵 (a_{ij}) 具有如下性质：

（1） (a_{ij}) 正定并且行列式为 1；

（2） 对角元 $a_{ij}(1\leqslant i\leqslant k)$ 均为偶数.

则必然 $k\equiv 0(\mathrm{mod}\ 8)$，并且 $\theta_f(\tau)\in M_{k/2}(\Gamma)$. 其证明和进一步例子可见 J. -P. Serre《数论教程》第七章.

以上我们介绍了模形式和它的数论背景. 19 世纪中期开始产生了模形式概念，具有了模形式的基本例子（爱森斯坦级数和 theta 函数）并且建立了模形式空间的基本事实. 模形式理论的下一个重大进步是由德国数学家 Hecke 于 20 世纪 30 年代做出的. 下面以模形式空间 $M_k(\Gamma_0(N))$ 为例来介绍 Hecke 理论.

Hecke 在模形式空间 $M_k(\Gamma_0(N))$ 中引入一批变换（现在叫作 Hecke **算子**）$T_k(m)(m=1,2,3,\cdots)$. 对于

$$f(\tau)=\sum_{n=0}^{\infty}c_nq^n\in M_k(\Gamma_0(N))\qquad(k\geqslant 1)$$

定义 $T_k(m)f=\sum_{n=0}^{\infty}b_nq^n$，其中对每个 $n\geqslant 0$，

$$b_n=\sum_{\substack{7\leqslant a\mid(n,m)\\(a,N)=1}}a^{k-1}c_{\frac{nm}{a^2}}.$$

例如 $b_1=c_m, b_0=c_0\sum_{\substack{7\leqslant a\mid m\\(a,N)=1}}a^{k-1}$. 而 $T_k(1)$ 是恒等映射. 要更好地理解这个定义和后面的重要事实，需要回到与模形式所对应的格上去. Hecke 算子在模形式 Fourier 系数上的作用是由这些算子在格上的自然作用诱导出来的. Hecke 证明了：若 $f\in M_k(\Gamma_0(N))$，则 $T_k(m)f\in M_k(\Gamma_0(N))$. 于是 $T_k(m)(m\geqslant 1)$ 是向量空间 $M_k(\Gamma_0(N))$ 中的一批线性算子. 然后他给出这些算子之间的如下关系：

定理 1 $M_k(\Gamma_0(N))$上的 Hecke 算子 $T_k(m)(m\geqslant1)$满足以下关系：

(1) 若$(m,n)=1$，则 $T_k(mn)=T_k(m)T_k(n)$（从而 $T_k(m)$和 $T_k(n)$是可交换的，并且把问题归结为 $T_k(p^r)$的情形）；

(2) 设 p 为素数，$r\geqslant1$，则当 $p\mid N$ 时，$T_k(p^r)=T_k(p)^r$. 而当 $P\nmid N$ 时，有如下递推公式

$$T_k(p^{r+1})=T_k(p^r)T_k(p)-p^{k-1}T_k(p^{r-1});$$

(3) 由(1)和(2)可知，所有的 Hecke 算子$\{T_k(m)\mid m\geqslant1\}$由$\{T_k(p)\mid p$ 为素数$\}$所决定，并且对任意 $n,m\geqslant1$，均有

$$T_k(n)T_k(m)=T_k(m)T_k(n).$$

如果我们形式地定义

$$H(s)=\sum_{n\geqslant1}T_k(n)n^{-s},$$

那么由(1)可知 $H(s)$有欧拉无穷乘积形式

$$H(s)=\prod_p H_p(s),H_p(s)=\sum_{r\geqslant0}T_k(p^r)p^{-rs}.$$

由(2)可知当 $p\mid N$ 时

$$H_p(s)=\sum_{r\geqslant0}T_k(p)^r p^{-rs}=(1-T_k(p)p^{-s})^{-1}$$

而由(3)中递推关系可知当 $p\nmid N$ 时

$$H_p(s)=(1-T_k(p)p^{-s}+p^{k-1-2s})^{-1}.$$

于是

$$\sum_{n\geqslant1}T_k(n)n^{-s}=\prod_{p\mid N}\frac{1}{1-T_k(p)p^{-s}}\cdot\prod_{p\nmid N}\frac{1}{1-T_k(p)p^{-s}+p^{k-1-2s}}.$$

下一个重要结果是：尖点模形式空间 $S_k(\Gamma_0(N))$可以引入一个内积(Petersson 内积)，使得 $S_k(\Gamma_0(N))$成为(有限维)希尔伯特空间. 而所有 Hecke 算子都是 $S_k(\Gamma_0(N))$到自身的线性算子，并

且均是此空间上的**自伴算子**. 由于这些 Hecke 算子彼此均可交换的. 根据希尔伯特空间理论, $S_k(\Gamma_0(N))$中一定存在一组基$\{f_i(\tau) \mid 1 \leqslant i \leqslant m\}$ $(m=\dim S_k(\Gamma_0(N)))$, 使得每个尖点模形式 $f_i(\tau)$都是所有 Hecke 算子 $T_k(n)(n \geqslant 1)$的本征函数, 即 $T_k(n) f_i(\tau)=\lambda_{k,i} f_i(\tau)$ $(\lambda_{k,i} \in \boldsymbol{C})$.

后人把是所有 Hecke 算子的本征函数的尖点模形式叫作 Hecke **模形式**. 设

$$f(\tau) = \sum_{n \geqslant 1} c_n q^n$$

是 $S_k(\Gamma_0(N))$中的 Hecke 模形式, 并且 $c_0=1$. 记 $T_k(m) f=\lambda_m f(m \geqslant 1)$. 由 Hecke 算子的定义知 $T_k(m) f$ 的 Fourier 展开首项系数为 c_m, 而 $\lambda_m f$ 的首项为 λ_m. 于是 $\lambda_m=c_m$, 即 f 对 $T_k(m)$的本征值 λ_m 就是 f 的 Fourier 系数 c_m. 由于当$(n, m)=1$ 时 $T_k(mn)=T_k(m) \cdot T_k(n)$. 将此等式两边作用于 f, 就得到: 当$(n, m)=1$ 时, $c_{nm}=c_n c_m$. 同样地, 当 $p \mid N$ 时, 由 $T_k(p^r)=T_k(p)^r$ 可知 $c_{p^r}=c_p^r$. 而当 $p \nmid N$ 时, 由 $T_k(p^{r+1})=T_k(p) T_k(p^r)-p^{k-1} T_k(p^{r-1})$ $(r \geqslant 1)$可知 $c_{p^{r+1}}=c_p c_{p^r}-p^{k-1} c_{p^{r-1}}$. 所以若令

$$L_f(s) = \sum_{n \geqslant 1} c_n n^{-s}$$

(这叫作**模形式** f **的** L-**函数**), 则它有欧拉无穷乘积展开

$$L_f(s) = \prod_{p \mid N} \frac{1}{1 - c_p p^{-s}} \cdot \prod_{p \nmid N} \frac{1}{1 - c_p p^{-s} + p^{k-1-2s}}. \quad (*)$$

可以证明: 对于每个权 k 的尖点模形式 $f=\sum_{n \geqslant 1} c_n q^n$, 对其 Fourier 系数均有估计 $c_n=O(n^{k/2})$(其中 O 是与 n 无关的). 由此可知级数 $L_f(s)$在 $\operatorname{Re}(s)>\frac{k}{2}+1$ 区域中收敛, 综合上述便有

定理 2 (1)$S_k(\Gamma_0(N))$中存在由 Hecke 模形式构成的一组基.

(2)设 $f=\sum_{n\geqslant 1}c_nq^n$ 是 $S_k(\Gamma_0(N))$中的 Hecke 模形式,并且 $c_0=1$,则 f 的 L-函数 $L_f(s)=\sum_{n\geqslant 1}c_nn^{-s}$ 在区域 $\mathrm{Re}(s)>\frac{k}{2}+1$ 中收敛,并且有形如(*)式的欧拉无穷乘积展开.

例 我们有爱森斯坦级数

$$E_4(\tau)=1+240\sum_{n\geqslant 1}\sigma_3(n)q^n=1+240q+\cdots\in M_4(\Gamma),$$

$$E_6(\tau)=1-504\sum_{n\geqslant 1}\sigma_5(n)q^n=1-504q+\cdots\in M_6(\Gamma).$$

于是

$$\Delta(\tau)=\frac{1}{1728}(E_4^3-E_6)=\sum_{n\geqslant 1}\tau(n)q^n\in S_{12}(\Gamma),\tau(1)=1.$$

从而 $\tau(n)$均为有理数.事实上,我们有另一个有趣的表达式

$$\Delta(\tau)=q\prod_{n\geqslant 1}(1-q^n)^{24},$$

所以 $\tau(n)$均是整数.注意 $\Gamma=\Gamma_0(1)$,而 $s_{12}(\Gamma)$是一维空间,按照 Hecke 理论它应当由 Hecke 模形式张成,所以 $\Delta(\tau)$本身就是 Hecke 模形式.于是 $\Delta(\tau)$的 L-函数为

$$L_\Delta(s)=\sum_{n\geqslant 1}\tau(n)q^n=\prod_p(1-\tau(p)p^{-s}+p^{11-2s})^{-1}.$$

对于 $L_\Delta(s)$的这个欧拉乘积展开是 1916 年由印度数学家 Ramanujan 提出的,在 1920 年由英国莫代尔(Mordell)所证明.由尖点模形式一般结果知 $\tau(p)=O(p^6)$(当 $p\to\infty$时),而 Ramanujan 猜想:对每个素数 p,$|\tau(p)|\leqslant 2p^{\frac{11}{2}}$.后来 Petersson 提出更一般的猜想:设 $f(\tau)\in S_{2k}(\Gamma)$,并且对某个素数 p,$T_{2k}(p)f=c_pf$,则 $|c_p|\leqslant 2p^{k-\frac{1}{2}}$.如果令

$$1-c_pp^{-s}+p^{2k-1-2s}=(1-\alpha p^{k-\frac{1}{2}-s})(1-\beta p^{k-\frac{1}{2}-s}),$$

则 $\alpha\beta=1$,$p^{\frac{k}{2}-1}(\alpha+\beta)=c_p$.而 Petersson 猜想等价于 $|\alpha|=|\beta|=$

1，或者等价于 α 和 β 是共轭的（注意 $T_{2k}(p)$ 是自伴算子，本征值 c_p 一定为实数）. 上述 Ramanujan－Petersson 猜想是有限域上高维代数簇的韦依猜想的一种推广，这两个猜想于 1973 年同时由 Deligne 证明.

例 现在我们用模形式理论寻求 $x_1^2+x_2^2+x_3^2+x_4^2=n$ 整数解个数 $N_4(n)$ 的一般公式. 我们知道

$$\theta^4(\tau) = \sum_{n\geqslant 0} N_4(n)q^n = 1 + 8q + 24q^2 + \cdots \in M_2(\Gamma_0(4)).$$

另一方面，可以证明

$$F(\tau) = \sum_{2\nmid n\geqslant 1} \sigma_1(n)q^n \in M_2(\Gamma_0(4)),$$

并且 θ^4 和 F 是 $M_2(\Gamma_0(4))$ 的一组基. 于是 $S_2(\Gamma_0(4))$ 是由 F 生成的一维空间，从而 F 是 Hecke 模形式. 特别地，F 是 Hecke 算子 $T_2(2)$ 的本征函数（以下 $T_2(2)$ 简记为 $T(2)$）. 由于 F 对 $T(2)$ 的本征值是 F 的 Fourier 展开中 q^2 的系数 0，从而 $T(2)F=0$. 但是由 $T(2)$ 的定义可知，若 $T(2)\left(\sum\limits_{n\geqslant 0} c_nq^n\right) = \sum\limits_{n\geqslant 0} b_nq^n$，则

$$b_n = \sum_{\substack{1\leqslant a\mid(n,2)\\(a,4)=1}} ac_{\frac{2n}{a^2}} = c_{2n}.$$

由此可知 $T(2)\theta^4 \neq 0$，即 $T(2)$ 在 $M_2(\Gamma_0(N))$ 上不恒为零，从而有非零本征值 λ. 设 $\theta^4+\beta F=\sum\limits_{n\geqslant 0} c_nq^n$ 是对应的本征函数，其中当 $2\mid n$ 时 $c_n=N_4(n)$，而 $2\nmid n$ 时 $c_n=N_4(n)+\beta\sigma_1(n)$. 则

$$\sum_{n\geqslant 0} N_4(2n)q^n = T(2)(\alpha\theta^4 + \beta F) = \lambda\sum_{n\geqslant 0} c_nq^n. \tag{1}$$

比较此式前两项系数可知

$$1 = N_4(0) = \lambda c_0 = \lambda N_4(0) = \lambda,$$
$$24 = N_4(2) = \lambda N_4(1) + \beta\sigma_1(1) = 8\lambda + \beta,$$

于是 $\lambda=1$，$\beta=16$，从而(1)式成为

$$\sum_{n\geqslant 0} N_4(2n)q^n = \sum_{n\geqslant 0} N_4(n)q^n + 16\sum_{2\nmid n} \sigma_1(n)q^n.$$

这表明：

当 $2\nmid n$ 时，$N_4(2n)=N_4(n)+\sigma_1(n)\cdot 16$； (2)

当 $2\mid n$ 时，$N_4(2n)=N_4(n)$. (3)

我们现在证明

$$\text{当 } 2\nmid n \text{ 时}, N_4(2n)=3N_4(n). \tag{4}$$

这是因为：方程 $x_1^2+x_2^2+x_3^2+x_4^2=2n$ 在 n 为奇数时必然 $\{x_1, x_2, x_3, x_4\}$ 是两奇两偶. 规定 x_1 和 x_2 为奇数，则这样的解数为 $\frac{N_4(2n)}{6}$.

令 $Y_1=\frac{x_1+x_2}{2}$，$Y_2=\frac{x_1-x_2}{2}$，$Y_3=\frac{x_3+x_4}{2}$，$Y_4=\frac{x_3-x_4}{2}$，则 $Y_1^2+Y_2^2+Y_3^2+Y_4^2=n$，并且 Y_1 和 Y_2 一奇一偶，而 Y_3 和 Y_4 同奇或同偶. 这样的解数为 $\frac{N_4(n)}{2}$. 满足上述条件的两个方程的整数解是一一对应的. 于是 $\frac{N_4(2n)}{6}=\frac{N_4(n)}{2}$，即(4)式成立.

由(2)式和(4)式知当 $2\nmid n$ 时，$N_4(n)=8\sigma_1(n)$. 再由(3)式知当 n 为偶数时，令 $n=2^r n_0$，其中 $r\geqslant 1$，$2\nmid n_0$，则 $N_4(n)=N_4(2n_0)=3N_4(n_0)=24\sigma_1(n_0)$.

现在介绍 Hecke 理论的最后一个重要结果，即模形式的 L-函数的解析延拓与函数方程. 和黎曼 zeta 函数 $\zeta(s)$ 的情形一样，解析延拓是用伽玛函数把模形式 $f(\tau)=\Sigma a_n e^{2\pi i\tau n}$ 和它的 L-函数 $L_f(s)=\Sigma a_n n^{-s}$ 联系起来，给出 $L_f(s)$ 的积分表示

$$\int_0^\infty f(i\sigma)\sigma^{s-1}\mathrm{d}\sigma=(2\pi)^{-s}\Gamma(s)L_f(s),$$

然后用模形式在模群作用下的变换公式得到 $L_f(s)$ 的函数方程.

我们仍以 $S_k(\Gamma_0(N))$ 为例. 在这个空间上可以定义一个对合算子 W_N（即 $W_N^2=I$），于是空间分解成直和

$$S_k(\Gamma_0(N))=S_k^+(\Gamma_0(N))\oplus S_k^-(\Gamma_0(N)).$$

其中对 $\varepsilon=\pm$，

$$S_k^{\varepsilon}(\Gamma_0(N))=\{f\in S_k(\Gamma_0(N))\,|\,W_N(f)=\varepsilon f\}.$$

定理 3 设 $f\in S_k^{\varepsilon}(\Gamma_0(N))$. 则在 $\mathrm{Re}(s)>\frac{k}{2}+1$ 中定义的级数 $L_f(s)$ 可解析延拓为整个复平面上的全纯函数，并且令 $\Lambda_f(s)=N^{s/2}(2\pi)^{-s}\Gamma(s)L_f(s)$，则有函数方程 $\Lambda_f(s)=\varepsilon(-1)^{k/2}\Lambda_f(k-s)$.

希尔伯特第 11 问题是问对于任意代数数域 K，研究二次型的求解问题，对于域 K 上求解问题就是 § 3.2 介绍的哈瑟整体-局部原则. 而对于整数环 O_K 上求解问题，则需要引进关于代数数域 K 的模形式，叫作希尔伯特模形式. 此外，若作高维推广，即研究二次型矩阵方程 $XAX^T=B$ 的整数解问题，则有 Siegel 模形式（在 $\mathbf{Z}$ 上）和 Hilbert-Siegel 模形式（在任意代数整数环 O_K 上）. 所有这些模形式理论的基本问题都是构作各种模形式空间的一组基，研究这些模空间的结构，半整权模形式的深入研究以及由它构作整权模形式（志村提升）. 模形式理论是数论、函数论、微分几何、代数几何、群表示论以及微分方程等相互融合的一个交叉学科，与它联系最密切的是在 20 世纪发展起来的另一个数论分支：椭圆曲线的算术理论.

参 考 文 献

[1] G. Eisenstein, Mathematische Abhandlungen, Berlin, 1847 (1967 年再版).

[2] E. Hecke, Mathematische Werke, Göttingen, 1959.

[3] S. Lang, Introduction to Modular Forms, Springer-Verlag, 1976.
[4] T. Miyake, Modular Forms, Springer-Verlag, 1989.
[5] R. Rankin, Modular Forms and Functions, Cambridge Univ. Press, 1977.
[6] J.-P. Serre, A Course in Arithmetic, GTM 7, Springer-Verslag, 1973(本书第二部分讲模形式).

§3.6 椭圆曲线的算术理论

代数几何考虑代数曲线，是把它们作双有理分类，不同亏格的代数曲线彼此不(双有理)等价. 当亏格≥1 时，甚至同样亏格的代数曲线也有许多等价类. 但是从数论观点，研究代数曲线在数域 K (或者整体域包括函数域)上的解，则所有代数曲线只分成三大类.

第一类是有理曲线，即亏格为 0 的曲线，包括直线和二次曲线. 定义在 K 上的有理曲线在 K 中或者无解，或者有无穷多个解. 第二类是亏格≥2 的曲线. 1923 年莫代尔猜想：有理数域 $\boldsymbol{Q}$ 上任何亏格≥2 的代数曲线只有有限多个有理数解. 1983 年法廷斯(Faltings)对于任意整体域的情形证明了这个猜想. 第三类就是亏格为 1 的代数曲线，就叫作椭圆曲线. 整体域上这种曲线可能有有限多解，也可能有无限多解. 更有趣的是：整体域上每个椭圆曲线的所有解可定义一个运算，使得这些解形成阿贝尔群，并且莫代尔证明了这个群是有限生成的. 这就引发出一连串的数论问题：如何计算这个群的秩?如何构作这个群的一组基(从而所有点都可由基中有限个点通过运算得到)?有限阶元素形成的有限子群有什么样

的结构？…

我们还是以有理数域作为例子．$\boldsymbol{Q}$ 上椭圆曲线 E 的典型方程为

$$E: y^2=x^3+ax+b, a, b\in \boldsymbol{Q}, \Delta=4a^3+27b^2 \neq 0.$$

其中 Δ 叫曲线 E 的判别式，$\Delta\neq 0$ 意味着三次多项式 x^3+ax+b 没有重根（在 $\boldsymbol{C}$ 中）．更一般地，域 K 上的椭圆曲线均双有理等价于 $y^2+exy+ay=x^3+bx^2+cx+d$，其中 $e,a,b,c,d\in k$．如果域的特征不为 2 或 3，则方程可以化成更简单形式：

$$y^2=x^3+ax+b,\quad a,b\in k, \Delta=4a^3+27b^2\neq 0.$$

古代人们就已经研究这类方程的有理数解（和整数解）．我们举两个例子．

例 1（同余数问题）　古希腊人研究如下的几何问题：哪些正整数 n 是三边均为有理数的直角三角形的面积．换句话说，是否存在正有理数 a,b,c，使得

$$a^2+b^2=c^2,\ \text{并且}\quad \frac{1}{2}ab=n. \tag{1}$$

如果(1)有解，称 n 为 Congruent number，我们故且叫作同余数．例如由 $(a,b,c)=(3,4,5)$ 可知 6 是同余数，而 1225 年意大利人 Fibonacci 发现 5 为同余数，因为可取 $(a,b,c)=\left(\frac{3}{2},\frac{20}{3},\frac{41}{6}\right)$．费马曾经证明了 1,2,3 均不是同余数．易知 n 为同余数当且仅当 nm^2 是同余数，于是 4 也不是同余数．今后我们总假设 n 是无平方因子的正整数．

公元 10 世纪左右，阿拉伯人研究另一个数论问题：对于正整数 n，是否

存在有理数 x，使得 $x, x+n, x-n$ 均为有理数平方．　　(2)

事实上，这两个问题是等价的：若 (a,b,c) 是问题(1)的正有理数解，则 $x=\frac{c^2}{4}$ 就是问题(2)的解．反之若 x 是问题(2)的有理数解，则 $c=2\sqrt{x}$，$a=\sqrt{x+n}+\sqrt{x-n}$，$b=\sqrt{x+n}-\sqrt{x-n}$ 就是问题

(1)的正有理数解.

如果 $x=D$ 是问题(2)的有理解，则 D 和 $D\pm n$ 均为正有理数 A,B,C 的平方. 所以椭圆曲线

$$E: y^2=x^3-n^2x$$

有有理数解 (x,y)，其中 $x=D, y=ABC\neq 0$. 反之，若 $(x,y)=(A,B)$ 是椭圆曲线 E 的有理数解并且 $B\neq 0$，可以直接验证 $\left(\frac{A^2+n^2}{2B}\right)^2$ 和 $\left(\frac{A^2+n^2}{2B}\right)^2\pm n$ 都是有理数平方，因为 $\left(\frac{A^2+n^2}{2B}\right)^2\pm n=\left(\frac{A^2\pm 2nA-n^2}{2B}\right)^2$. 所以 n 是否为同余数又等价于是否

$$\text{椭圆曲线 } E \text{ 存在有理数解 } (x,y) \text{ 使得 } y\neq 0. \tag{3}$$

事实上，利用我们将要定义的加法运算，由椭圆曲线 E 的一组有理数解 $(x,y)(y\neq 0)$ 可得到 E 的无穷多组有理数解. 所以 n 为同余数当且仅当 E 有无穷多组有理数解. 利用椭圆曲线的深刻理论，目前已经决定出许多类同余数和非同余数. 但这个问题至今没有完全解决. 比如说：1960 年有三位数学家猜想：每个正整数 $n\equiv 5,6,7 \pmod 8$ 都是同余数. 这个猜想至今也未能证明. 我们在后面还会回到这个问题.

例 2 丢番都在《算术》一书中研究椭圆曲线

$$E: y(6-y)=x^3-x$$

的有理数解. 除了使方程两边均为 0 的平凡解 $x=0,\pm 1$ 和 $y=0, 6$ 之外，他还得到新的解：令 $x=3y-1$，代入 E 的方程则为

$$6y-y^2=x(x+1)(x-1)=3y(9y^2-9y+2),$$

即 $y^2(27y^2-26y)=0$. 此方程有根 $y=\frac{26}{27}$，而 $x=3y-1=\frac{17}{9}$，于是得到解 $(x,y)=\left(\frac{17}{9},\frac{26}{27}\right)$.

为什么要用代换 $x=3y-1$？丢番都并没有任何解释. 据说这个解释是由伟大的英国数学家和物理学家牛顿给出的(见 A. W. Knapp:《Elliptic Curves》一书第 10 页). 假如我们知道椭圆曲线

$$E: y^2+axy+by$$
$$=cx^3+dx^2+ex+f \quad (a,b,c,d,e,f\in \boldsymbol{Q}\ c\neq 0)$$

上的两个不同的有理点 $A=(x_1,y_1)$ 和 $B=(x_2,y_2)$. 过这两点的直线 L 有方程 $\alpha x+\beta y=\gamma$,其中 $\alpha,\beta,\gamma\in\boldsymbol{Q}$,并且 α 和 β 不全为零. 不妨设 $\alpha\neq 0$,将 $x=-\dfrac{\beta}{\alpha}y+\dfrac{\gamma}{\alpha}$ 代入 E 的方程之后,便得到关于 x 的一个三次方程 $f(y)=0$. 其中 $f(y)$ 是系数为有理数的三次多项式. 由于点 A 和点 B 是椭圆曲线 E 和直线 L 的交点,可知有理数 y_1 和 y_2 都是 $f(y)$ 的根,但是 $f(y)$ 的系数也都是有理数. 所以 $f(y)$ 的第三个根 y_3 也是有理数. 令 $x_3=-\dfrac{\beta}{\alpha}y_3+\dfrac{\gamma}{\alpha}$,易知 (x_3,y_3) 是椭圆曲线 E 的有理点,它就是直线 L 和椭圆曲线 E 的第三个交点. 如果我们把点 B 沿着椭圆曲线移到点 A,那么直线 L 就变成椭圆 E 过点 A 的切线,所以即使我们只知道 E 上一个有理点 A,过此点作 E 的切线,它与 E 相交的另一个点也是 E 的有理点. 牛顿说,丢番都所用的直线 $x=3y-1$ 就是椭圆曲线 $y(6-y)=x^3-x$ 在其有理点 $(x,y)=(-1,0)$ 处的切线. 这条切线与椭圆曲线交于三点,其中 $(-1,0)$ 是二重点,另一个交点就是丢番都给出的新有理解 $(x,y)=\left(\dfrac{17}{9},\dfrac{26}{27}\right)$(见上图).

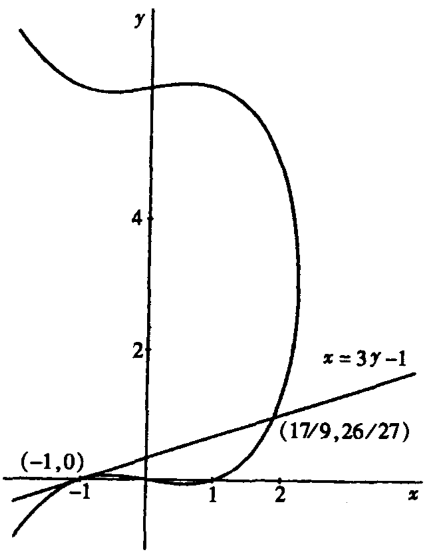


图 13　牛顿对丢番都方法的解释

牛顿的这种解释给出椭圆曲线上有理点之间的一种运算. 现

在用方程

$$E: y^2 = x^3 + ax + b \quad (a, b \in \boldsymbol{Q}, 4a^3 + 27b^2 \neq 0)$$

来定义这种运算. 若$(x,y)=(a,b)$为 E 上点,则$(a,-b)$也为 E 上一点,从而曲线 E 关于 x 轴是对称的. 现在设 $P=(x_1,y_1)$和 $Q=(x_2,y_2)$是椭圆曲线 E 上两个不同的有理点. 如上所述,联接点 P 和 Q 的直线 l 与曲线 E 的第三个交点(x_3,y_3)也是 E 上的有理点. 我们定义它的对称点 $R=(x_3,-y_3)$为点 P 和 Q 之和,即 $P\oplus Q=R$(见下图).

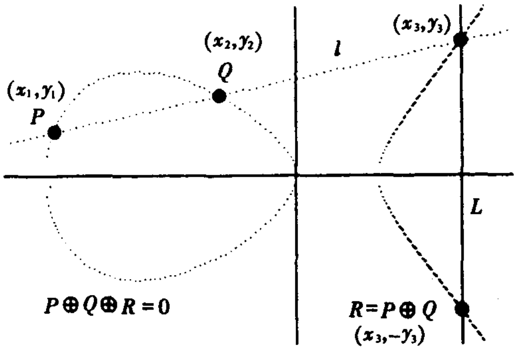


图 14 椭圆曲线上的加法运算

现在我们给出这种运算的解析表达式.

先设 $x_1 \neq x_2$. 则联接点 P 和 Q 的直线有方程

$$L: y-y_1 = m(x-x_1), m=\frac{y_2-y_1}{x_2-x_1} \in \boldsymbol{Q}.$$

将 $y=y_1+m(x-x_1)$代入 E 的方程得到

$$\begin{aligned} 0 &= (y_1+mx-mx_1)^2 + x^3 + ax + b \\ &= x^3 - m^2x^2 + \alpha x + \beta, \alpha, \beta \in \boldsymbol{Q}. \end{aligned}$$

由于 x_1 和 x_2 是此方程两个根,所以另一个根为

$$x_3 = -x_1 - x_2 + m^2 \in \boldsymbol{Q},$$

而 $y_3=y_1+m(x_3-x_1)\in\boldsymbol{Q}$. 于是 $P\oplus Q=R$,其中 $R=(x_3,-y_3)$.

如果 $x_1=x_2$,而 P 和 Q 是两个不同的有理点,必然 $y_1=-y_2$

$\neq 0$. 这时由上面公式算出 $m=\infty$. 几何上,联接 $P=(x_1,y_1)$和 $Q=(x_1,-y_1)(y_1\neq 0)$的直线 L 是与 y 轴平行的直线,它与曲线 E 只有两个交点 P 和 Q. 于是我们要在椭圆曲线上增加一个"无穷远点"∞,它是与 y 轴平行的所有直线的公共交点. 设想把∞和 $P=(x_1,y_1)$相加,联结∞和 P 的直线就是过 P 与 y 轴平行的直线,它与 E 的第三交点为$(x_1,-y_1)$,而$(x_1,-y_1)$的对称点就是 $P=(x_1,y_1)$. 因此$\infty\oplus P=P$,即∞是这种运算的零元素,所以今后把无穷远点∞改记成 O. 记 $\overline{P}=(x_1,-y_1)$,则 $P\oplus\overline{P}=\overline{O}=O$,于是 $\overline{P}=-P$.

我们再考虑 $P=Q=(x_1,y_1)$的情形,令 l 为曲线 E 过点 P 的切线,它与 E 的交点(x_3,y_3)可由简单的解析几何算出:当 $y_1\neq 0$ 时

$$\begin{cases} x_3=-2x_1+m^2, \\ y_3=y_1+m(x_3-x_1), m=\dfrac{3x_1^2+a}{2y_1}, \end{cases}$$

即$[2]P=P\oplus P=R, R=(x_3,-y_3)$. 而当 $y_1=0$ 时,$[2]P=O$.

以上给出了椭圆曲线有理点(包括 0)加法运算的完全定义. 由几何或解析表达式都容易看出运算满足交换律,但结合律却并不显然. 用解析表达式可以验证结合律(尽管计算相当复杂). 于是,椭圆曲线的所有有理点(连同 O)一起对上述加法构成阿贝尔群. 对于任意域 K,定义在 K 上的椭圆曲线 E,它的坐标属于 K 的所有点连同 O 都可类似地作成一个阿贝尔群,记此群为 $E(K)$.

例 3 考虑方程 $u^3+v^3=1$. 由

$$\begin{cases} u=\dfrac{3x}{y}, \\ v=\dfrac{y-9}{y}, \end{cases} \qquad \begin{cases} x=\dfrac{3u}{1-v}, \\ y=\dfrac{9}{1-v}, \end{cases}$$

可知 $u^3+v^3=1$ 与椭圆 $y^2-9y=x^3-27$ 双有理等价. 于是曲线 $u^3+v^3=1$ 也是椭圆曲线,即亏格为 1(一般地,可证 $u^n+v^n=1$ 的亏

格为$\frac{1}{2}(n-1)(n-2)$). 欧拉证明了 $u^3+v^3=w^3$ 只有平凡整数解，由此可知 $u^3+v^3=1$ 只有两组有理数解$(u,v)=(1,0),(0,1)$. 再由上述双有理变换，可知 $y^2-9y=x^3-27$ 只有有理数解$(x,y)=(3,9),(3,0)$. 由于$\left(y-\frac{9}{2}\right)^2=y^2-9y+\frac{81}{4}=x^3-27+\frac{81}{4}=x^3-\frac{27}{4}$，可知对椭圆曲线 $E:y^2=x^3-\frac{27}{4}$，$E(Q)=\left\{\left(3,\pm\frac{9}{2}\right),0\right\}$. 从而 $E(Q)$是三阶循环群. 直接用加法公式也可证$[2]\left(3,\frac{9}{2}\right)=\left(3,\frac{9}{2}\right)\oplus\left(3,\frac{9}{2}\right)=\left(3,-\frac{9}{2}\right)$，从而$[3]\left(3,\frac{9}{2}\right)=\left(3,\frac{9}{2}\right)\oplus\left(3,-\frac{9}{2}\right)=O$.

对于椭圆曲线 $E:y^2=x^3-n^2x$(n 为无平方因子正整数)，$(0,0),(\pm n,0)$都是群 $E(Q)$的二阶元素. 可以证明它们和 O 是$E(Q)$的全部有限阶元素. 于是若 $E(Q)$还有点$(x,y)(y\neq 0)$，则必为无限阶元素，即 $E(Q)$必为无限阿贝尔群(我们在前面知道，这相当于 n 是同余数).

椭圆曲线的下一个重要结果是莫代尔于 1923 年证明的：Q 上的任意一条椭圆曲线 E，$E(Q)$均是有限生成阿贝尔群. 他的方法是：对于 $E(Q)$中每个点引入一个“高度”概念来衡量此点的“大小”. 记 $nE(Q)=\{[n]P:P\in E(Q)\}$，这是 $E(Q)$的子群. 先证 $E(Q)$的有限阶元素只有有限多个，再对于有理点的高度采用费马的“无穷下降”方法证明 $E(Q)/2E(Q)$是有限群，由此即知 $E(Q)$是有限生成的. 用此方法可以证明更一般的结果：对于任意整体域 K(数域或函数域)，E 为 K 上任意椭圆曲线，则 $E(K)$均是有限生成阿贝尔群，即

$$E(k)=E(k)_t\oplus E(K)_f \quad (\text{直和}),$$

其中 $E(k)_t$ 是 $E(k)$的所有有限阶元素构成的有限阿贝尔群，而 $E(K)_f$是秩 r 的自由阿贝尔群，其中 $r=r(E)$是非负整数，叫作椭

圆曲线 E 的秩，于是：$E(K)$有限当且仅当 $r(E)=0$. 对每个具体的椭圆曲线 E，决定有限子群 $E(K)_t$ 通常是容易的. 所以椭圆曲线算术理论的基本问题是：决定 $E(K)$的秩 r 并且给出 $E(K)_f$ 的一组基$\{P_i|1\leqslant i\leqslant r\}$. 于是 $E(K)$中每个元素均可惟一表示成

$$P=P'\oplus n_1P_1\oplus\cdots\oplus n_rP_r \quad (n_i\in\boldsymbol{Z},P'\in E(K)_t),$$

即由有限阶元素和基中元素（共有限个）通过加法运算$\oplus$可得到 $E(K)$的全部元素.

关于群 $E(K)$的结构有如下两个重要猜想：对每个固定的代数数域 K，

（1） K 上所有椭圆曲线 E 的有限阿贝尔群 $E(K)_t$ 只有有限多种可能的结构.

（2） 对每个正整数 N 均存在 K 上椭圆曲线 E，使得群 $E(K)$ 的秩大于 N.

1978 年，美国哈佛大学梅祖尔（B. Mazur）对于 $K=\boldsymbol{Q}$ 的情形首先解决了问题（1）：$E(\boldsymbol{Q})_t$ 只有以下 15 种可能

$$\boldsymbol{Z}_n(n=1,2,\cdots,9,10,12),\boldsymbol{Z}_{2n}\oplus\boldsymbol{Z}_2(n=1,2,3,4).$$

1997 年，问题（1）对任意数域 K 都被证明. 至于问题（2），目前对 $K=\boldsymbol{Q}$ 也未解决，只构作出 $\boldsymbol{Q}$ 上椭圆曲线 E，使得 $E(\boldsymbol{Q})$的秩$\geqslant 14$.

20 世纪中期，椭圆曲线算术理论另一个重大进步是解析数论的介入. 就像用狄德金 zeta 函数 $\zeta_k(s)$的解析特性（在 $s=1$ 的留数）来反映数域 K 的代数和算术特性（理想类数，单位群等）一样，人们希望由数域 K 上椭圆曲线 E 构作一个恰当的解析函数，使这个函数的解析特性能反映群 $E(K)$的代数结构. 我们仍以 $\boldsymbol{Q}$ 为例. 设椭圆曲线为

$$E: y^2=x^3+ax+b,\ a,\ b\in\boldsymbol{Z},\ \Delta=4a^3+27b^2\neq 0.$$

如果 p 为素数并且 $p\nmid\Delta$，则在有限域 $\boldsymbol{F}_p=\boldsymbol{Z}/p\boldsymbol{Z}$ 中 $\Delta\neq 0$，所以 E 在 $\boldsymbol{F}_p$ 上也是一条椭圆曲线，把它记成 E_p. E_p 在有限域 $\boldsymbol{F}_p$ 中的点数是有限的，记为 $N_p=N_p(E_p)=|E_p(\boldsymbol{F}_p)|$. 满足 $p\nmid\Delta$ 的素数 p 有无穷多个，人们希望这无穷多个 N_p 值能给出有理点群 $E(\boldsymbol{Q})$

的结构，即希望用无穷多个有限域 $\boldsymbol{F}_p(p \nmid \Delta)$ 上椭圆曲线 E_p 的韦依 zeta 函数来构作椭圆曲线 E 的一个整体的复变函数，使得它的解析性质反映 $E(\boldsymbol{Q})$ 的结构.

根据 § 3.4 所述，当 $p \nmid \Delta$ 时，$\boldsymbol{F}_p$ 上椭圆曲线 E_p 的 zeta 函数为（注意曲线 E_p 的亏格为 1）

$$\zeta_{E_p}(s)=\frac{L(p^{-s})}{(1-p^{-s})(1-p^{1-s})},$$

其中

$$L(p^{-s})=1-a_p p^{-s}+p\cdot p^{-2s}=(1-w_p p^{-s})(1-\overline{w}_p p^{-s}),$$

而系数 a_p 与 N_p 的关系为

$$N_p=1+p-w_p-\overline{w}_p=1+p-a_p.$$

于是 $\zeta_{E_p}(s)$ 由 a_p 所决定，从而也由 N_p 所决定. 20 世纪 50 年代人们用这些 zeta 函数的分子 $L(p^{-s})$ 构作出椭圆曲线 E 的一个解析函数

$$L(E,s)=\prod_{p\mid\Delta}L_p(p^{-s})\cdot\prod_{p\nmid\Delta}(1-a_p p^{-s}+p^{1-2s})^{-1}\quad(1)$$

叫作椭圆曲线 E 的 L-函数，其中对满足 $p\mid\Delta$ 的有限个素数 p，局部因子 $L_p(p^{-s})$ 是关于 p^{-s} 的某个简单多项式的倒数，具体定义涉及到曲线 E 模 p 之后的奇点精细几何结构，此处从略. 由于韦依定理：$|a_p|=|w_p+\overline{w}_p|\leqslant 2\sqrt{p}$，可知定义 $L(E,s)$ 的级数在 $\mathrm{Re}(s)>\frac{3}{2}$ 时收敛. 1953 年～1957 年，德国人 Deuring 对一类特殊的"有复乘"椭圆曲线研究了它们的 L-函数，成功地把这些 L-函数解析延拓成整个复平面上的解析函数，并且满足函数方程.

对于 $\boldsymbol{Q}$ 上的每个椭圆曲线 E，有理点群 $E(\boldsymbol{Q})$ 是有限生成阿贝尔群，即是秩有限的 $\boldsymbol{Z}$-模. 整数环 $\boldsymbol{Z}$ 自然地作用在群 $E(\boldsymbol{Q})$ 上，即对每个正整数 n 和点 $P\in E(\boldsymbol{Q})$，$[n]P$ 为 n 个 P（对运算 $\oplus$）相加，而 $[-n]P=-[n]P$. 事实上，对某些椭圆曲线 E，如果考虑 E 在复数域上所有点组成的集合 $E(\boldsymbol{C})$，它对于前面定义的加法运算 $\oplus$ 也形成交换群. 这个群解析同构于加法群 $\boldsymbol{C}/L$，其中 L 是 $\boldsymbol{C}$ 的

某个格. 例如在研究同余数问题时出现的椭圆曲线 $E_n: y^2=x^3-n^2x$,它对应的格为 $L=w_N(\mathbf{Z}+i\mathbf{Z})=w_N\mathbf{Z}[i]$, 其中 w_N 是与 N 有关的一个复数. 由于虚二次域 $K=\mathbf{Q}(i)$的整数环 $\mathbf{Z}[i]$自然作用于格 $L=w_N\mathbf{Z}[i]$上, 由此可诱导出 $\mathbf{Z}[i]$在椭圆曲线复点群 $E(\mathbf{C})$($\cong \mathbf{C}/L$)上的作用. 具体说来,如果$(x,y)\in E(\mathbf{C})$, 易知$(-x,iy)\in E(\mathbf{C})$. 把 $E(\mathbf{C})$中的这个映射记为 $\varphi(x,y)=(-x,iy)$,则 $\varphi^2(x,y)=\varphi(-x,iy)=(-(-x),i(iy))=(x,-y)=[-1](x,y)$. 于是 $\varphi^2=[-1]$, 这表明 φ 相当于 i 在 $E(\mathbf{C})$上的作用,即对每个 $P\in E(\mathbf{C})$,$\varphi(P)=[i]P$. 而对 $\mathbf{Z}[i]$中每个高斯整数 $a+bi(a,b\in\mathbf{Z})$,$[a+bi]P=[a]P\oplus[b][i]P$. 所以群 $E(\mathbf{C})$允许有比 $\mathbf{Z}$ 大的"乘法", 即它是 $\mathbf{Z}[i]$-模. 不难证明:$\mathbf{C}$ 的每个格 L 的自同构群或者为 $\mathbf{Z}$,或者是某个虚二次域 K 的整数环 O_K,于是群 $E(\mathbf{C})$的乘法或者只有 $\mathbf{Z}$,或者允许扩大成 O_K. 对于后一种情形,我们称椭圆曲线 E 有"复乘". 由于虚二次域的整数环可作为 $\mathbf{C}$ 上某个格 L 的自同构群,利用以 L 为周期格的椭圆函数作出对应的椭圆曲线 E,使得 $E(\mathbf{C})\cong\mathbf{C}/L$. 再利用椭圆曲线理论和类域论便由椭圆函数的特殊点值构作出虚二次域的最大阿贝尔扩张. 到了 20 世纪 50 年代, Deuring 对于有复乘 O_K 的椭圆曲线 E,计算出 E 的 L-函数恰好相当于虚二次域 k 的一种 L-函数 $L(s,\chi_K)$,叫作 K 的 Hecke L-函数. Hecke 于 20 世纪 30 年代研究了这类函数可以解析延拓到整个复平面上,并且有函数方程. 于是对于具有复乘的椭圆曲线 E,Deuring 把 E 的 L-函数 $L(E,s)$解析延拓成整个复平面上的全纯函数(没有极点),并且有如下形式的函数方程:存在正整数 N,使得对

$$\wedge(s)=\left(\frac{\sqrt{N}}{2\pi}\right)^s\Gamma(s)L(E,s)$$

有 $\wedge(S)=\varepsilon\wedge(2-s)$,其中 $\varepsilon=\pm1$. (2)

比如对椭圆曲线 $E_n: y^2=x^3-n^2x$,可取

$$N=\begin{cases}32n^2, 若\ 2\nmid n,\\ 16n^2, 若\ 2\mid n,\end{cases}\qquad \varepsilon=\begin{cases}1, 若\ n\equiv1,2,3 \pmod 8,\\ -1, 若\ n\equiv5,6,7 \pmod 8).\end{cases}$$

基于大量没有复乘的椭圆曲线的计算例子，英国两位数学家 Birch 和 Swinnerton-Dyer 于 1960 年～1965 年关于 $L(E,S)$提出一系列猜想. 第一个便是

(BSD Ⅰ) 对于 $\boldsymbol{Q}$ 上所有椭圆曲线 E，它的 L-函数均可解析延拓成整个复平面上的全纯函数，并且满足形如(2)式的函数方程.

如果 $L(E,s)$是全纯函数，进而就要问：它的解析特性(在某些点的特殊值，零点和它们的阶等)是否反映群 $E(\boldsymbol{Q})$的代数结构. 定义 $L(E,s)$的级数(1)在 $\mathrm{Re}(s)>3/2$ 时是收敛的. 如果我们形式地把 $s=1$ 代入，撇开 $p|\Delta$ 那个有限乘积部分，则

$$\begin{aligned} L(E,1) &\sim \prod_{p\nmid\Delta}(1-a_p p^{-s}+p^{1-2s})^{-1}\Big|_{s=1} \\ &\sim \prod_{p\nmid\Delta}\left(1-\frac{a_p}{p}+\frac{1}{p}\right)^{-1}=\prod_{p\nmid\Delta}\frac{p}{p+1-a_p} \\ &=\prod_{p\nmid\Delta}\frac{p}{N_p}, \end{aligned}$$

其中 N_p 表示 $\boldsymbol{F}_p$ 上椭圆曲线 $E_p=E(\bmod p)$在 $\boldsymbol{F}_p$ 中的点数(包括无穷远点). 上式右边收敛指的是无穷乘积的极限为正实数. 如果对几乎所有 $p\nmid\Delta$，曲线 E_p 在有限域 $\boldsymbol{F}_p$ 中点数都很多，即 N_p 值均很大，则使无穷乘积为 0，即 $s=1$ 为 $L(E,s)$的零点，人们期望 $E(\boldsymbol{Q})$也很大，比如说 $E(\boldsymbol{Q})$是无限集合，即群 $E(\boldsymbol{Q})$的秩 $r_E\geqslant 1$. 更进一步，人们甚至期望 $L(E,s)$在 $s=1$ 处的零点阶等于群$E(\boldsymbol{Q})$的秩. 上述二位英国人基于大量例子提出了这个猜想，即

(BSD Ⅱ)对每个 $\boldsymbol{Q}$ 上的椭圆曲线 E，其 L-函数 $L(E,s)$在 $s=1$处的零点阶 R_E 等于有理点群 $E(\boldsymbol{Q})$的秩 r_E.

再进一步，如果 $R_E=0$，即 $s=1$ 不是 $L(E,s)$的零点，则 $L(E,1)\neq 0$. 按照猜想(BSD Ⅱ)，应当 $r_E=0$，即 $E(\boldsymbol{Q})$是有限群.

他们的下一个猜想是函数值 $L(E,1)$ 反映有限阿贝尔群 $E(\boldsymbol{Q})$ 的阶和群 $E(\boldsymbol{Q})$ 的其他特性.

(BSD Ⅲ)若 $L(E,1)\neq 0$,则

$$L(E,1)=\frac{A|Ш(E)|}{|E(\boldsymbol{Q})|^2},$$

其中 A 是与 E 的几何特性有关的某些不变量(实周期,玉河数等)的乘积,而 $Ш(E)$ 是 E 的所谓沙发列维奇群. 后者反映椭圆曲线不符合局部-整体原则的程度,猜想它是有限群.

以上三条便是关于椭圆曲线的 Birch 和 Swinnerton-Dyer 猜想,简称为 BSD 猜想. 对于任意代数数域(甚至函数域)可以提出类似的猜想. 它是椭圆曲线算术理论的重要猜想. 自 20 世纪 50 年代 Deuring 对有复乘的椭圆曲线证明了(BSD Ⅰ)之外,20 年后才有理论上的突破. 1977 年,J. Coates 和 A. Wiles 证明了:对于有复乘的椭圆曲线 E,若 $r_E\geqslant 1$(即 $E(\boldsymbol{Q})$ 为无限群),则 $R_E\geqslant 1$(即 $L(E,1)=0$). 特别地由 $L(E,1)\neq 0$ 可推出 $E(\boldsymbol{Q})$ 是有限群. 1983 年,B. Gross 和 D. Zagier 对于 $\boldsymbol{Q}$ 上另一类"模"椭圆曲线 E(可由模函数参数化的椭圆曲线)证明了:若 $R_E=1$,则 $r_E\geqslant 1$. 1988 年前苏联数学家 Kolyvagin 对于"模"椭圆曲线证明了 $Ш(E)$ 是有限群. 1989 年—1990 年,K. Rubin 对于 $\boldsymbol{Q}$ 上有复乘的椭圆曲线 E,证明了 $Ш(E)$ 是有限群,并且当 $L(E,1)\neq 0$ 时,如果把猜想(BSD Ⅲ)中等式改写成 $\frac{|E(\boldsymbol{Q})|^2}{A}L(E,1)=|Ш(E)|$,Rubin 证明了此式两边均是非零有理整数(的平方),并且两边的奇数部分是一致的. 对于具有复乘的许多椭圆曲线也证明了偶数部分的一致性,但没有完全解决. 20 世纪 70 年代以来上述一系列年青数学家关于 BSD 猜想的突破性结果,反映在这个时期椭圆曲线的算术理论不断取得进步,而这些结果的取得则采用了近代代数数论的各种新思想和新方法,包括半整权模形式理论和在下节讲述的近代分圆域理论

中构作 p-adic L-函数方法.

让我们举例说明 BSD 猜想对同余数问题的影响. 对于无平方因子正整数 n,椭圆曲线 $E_n: y^2=x^3-n^2x$ 是有复乘 $\mathbf{Z}[i]$的曲线. 我们已经知道:

n 为同余数$\Leftrightarrow|E_n(Q)|$无限$\Leftrightarrow r_{E_n}\geqslant 1$.

另一方面,Deuring 已经把 $L(E_n,s)$解析延拓成整个复平面上的全纯函数,即(BSD Ⅰ)对 E_n 成立. 而由 Coates-Wiles 结果,若 $L(E_n,1)\neq 0$,则$|E_n(Q)|$有限,从而 n 是非同余数,并且 Rubin 证明了猜想等式

$$\frac{|E_n(Q)|^2}{A}L(E_n,1)=|\text{Ш}(E_n)| \tag{3}$$

的两边都是非零有理整数的平方,并且奇数部分已证相等. 利用权为 3/2 的半整权模形式理论,Tunnel 对(3)式左边计算出一个非常初等的表达式:以 $N\{n=ax^2+by^2+cz^2\}$表示方程 $n=ax^2+by^2+cz^2$ 的整数解个数(其中 a,b,c 均为正整数). 则

(3) 式左边=

$$\begin{cases} N\{n=2x^2+y^2+32z^2\}-\frac{1}{2}N\{n=2x^2+y^2+8z^2\}, & \text{若 } 2\nmid n, \\ N\left\{\frac{n}{2}=4x^2+y^2+32z^2\right\}-\frac{1}{2}N\left\{\frac{n}{2}=4x^2+y^2+8z^2\right\}, & \text{若 } 2\mid n. \end{cases}$$

若 n 为同余数,则 $L(E_n,1)=0$,从而上式右边为 0. 反之若上式右边为 0,则 $L(E_n,1)=0$,即 $R_{E_n}\geqslant 1$. 由猜想(BSD Ⅰ)推出 $r_{E_n}\geqslant 1$,于是$|E_n(Q)|$无限,即 n 为同余数. 综合上述便得到:

定理(Tunnel,1983) 设 n 为无平方因子正整数. 如果 n 是同余数,则

$$N\{n=2x^2+y^2+32z^2\}=\frac{1}{2}N\{n=2x^2+y^2+8z^2\},\text{若 } 2\nmid n,$$

$$N\left\{\frac{n}{2}=4x^2+y^2+32z^2\right\}=\frac{1}{2}N\left\{\frac{n}{2}=4x^2+y^2+8z^2\right\},\text{若 } 2\mid n.$$

反之，若上面等式成立，则在(BSD Ⅱ)猜想之下 n 为同余数.

现在我们介绍椭圆曲线的另一个重要的猜想. 将定义 $L(E,s)$ 的欧拉无穷乘积公式(1)和 Deuring 对有复乘椭圆曲线的函数方程(2)与上节定理 2 中 Hecke 模形式的无穷乘积展开和定理 3 中的函数方程相比，可以看到它们完全一样($k=2$). 于是便有如下的

谷山-志村-韦依猜想(Taniyama-Shimura-Weil 猜想，简称 *TSW* 猜想)　对于 $\boldsymbol{Q}$ 上每条椭圆曲线 E，均存在正整数 N，使得 E 的 L-函数 $L(E,s)$ 是尖点模形式空间 $S_2(\Gamma_0(N))$ 中某个 Hecke 模形式 $f=\sum_{n\geq1}a_nq^n$ 的 L- 函数 $L_f(s)=\sum_{n\geq1}a_nn^{-s}$.

1955 年 9 月，谷山(1927—1958)在东京的一个国际学术会议上向人们展示了他计算的一些例子，表明椭圆曲线 L 函数展开的系数和某些模形式 Fourier 展开系数惊人地一致. 多数人认为这只是偶然的巧合，而志村五郎是支持谷山观点的少数人当中的一个，即相信所有椭圆曲线的 L-函数都来源于某种模形式. 在年青的谷山不幸去世之后，志村继续从代数几何角度研究这个猜想的本质，甚至认为"人们可以把一个通过模形式参数化的椭圆曲线看作是善和美的椭圆曲线. 我期望所有的椭圆曲线都是善和美的."梅祖尔在回忆这个猜想早期的情形时说："这是一个神奇的猜想，推测每个椭圆曲线相伴着一个模形式，但是一开始它就被人们忽视了，因为它太超前于它的时代."这个猜想所提出的问题，是要判定一个狄里赫利级数 $\sum_{n\geq1}a_nn^{-s}$ 需要满足什么样的条件，才能使级数 $\sum_{n\geq1}a_nq^n(q=e^{2\pi i\tau})$ 是模形式. Hecke 在 20 世纪 30 年代就研究过这个问题. 在模形式和椭圆曲线理论经过 30 年长足发展之后，韦依于 1967 年又进一步研究这个问题，并且以更明确形式再次提出

上述猜想.人们已经从几何和群表示论的角度发现了 TSW 猜想的其他等价形式.我们在前面只给出它最容易陈述的一种形式.

这个猜想揭示出椭圆曲线和模形式之间的深刻联系,它有许多重要的推论.比如说,我们在上节讲到,Hecke 在 20 世纪 30 年代就证明了所有 Hecke 模形式的 L-函数可以解析延拓并有函数方程.如果 TSW 猜想成立,即椭圆曲线的 L-函数都是 Hecke 模形式的 L-函数,那么椭圆曲线的 L-函数也可解析延拓并有同样的函数方程.也就是说:由 TSW 猜想可推出猜想(BSD Ⅰ).最惊人的消息发生在 1986 年:由 TSW 猜想可推出费马猜想.正是这个消息使模形式和椭圆曲线算术理论的顶尖高手 A. 怀尔斯面壁 7 年,对于一类重要的椭圆曲线证明了 TSW 猜想,并由此证明了具有 350 年历史的费马猜想.到了 2000 年,TSW 猜想已完全被证明.

参 考 文 献

[1] D. Husemöller, Elliptic Curves, GTM 111, Springer-Verlag, 1987.

[2] A. W. Knapp, Elliptic Curves, Princeton Univ. Press, 1992.

[3] N. Koblitz, Introduction to Elliptic Curves and Modular Forms, GTM 97, Springer-Verlag, 1984.

[4] S. Lang, Elliptic Functions, Springer-Verlag(第二版)1987.

[5] J. H. Silverman, The Arithmetic of Elliptic Curves, GTM 106, Springer-Verlag, 1986.

§ 3.7 近代分圆域理论(Ⅰ):p-adic L-函数

19 世纪中期,库默尔在研究费马猜想的过程中,对分圆域的整数环,理想类数和单位群作了奠基性工作. 20 世纪前半期人们对分圆域的研究主要是对阿贝尔域的类群结构和类数公式作细致分析,代表和总结性工作是 Hasse 于 1952 年所写的《阿贝尔域类数》一书. 从 50 年代末期开始,分圆域研究中逐渐出现了一系列新的方法,得到一批新的结果,也提出新的重要猜想. 在近代分圆域理论中使用 p-adic 分析构作了 p-adic L-函数和 zeta 函数,用有限群表示理论研究类群和单位群的伽罗华模结构,研究类群和单位群在一系列数域扩张($\boldsymbol{Z}_p$-扩张)中的性状,以及 p-adic 分布和测度理论. 这些在近代分圆域研究中所使用的工具和研究方法后来对数论其他分支(特别是模形式和椭圆曲线理论)产生了很大的影响. 在本节中我们重点介绍构作 p-adic zeta 函数和 L-函数的 p-adic 分析方法,下节介绍有限群表示理论在分圆域研究中的应用.

Hensel 于 20 世纪初建立了 p-adic 数和赋值理论,导致局部域理论、局部-整体原则和数论与代数几何的融合. 另一方面,p-adic 分析工具也象实分析那样沿着极限,连续,微分,积分的途径着手建立. 首先需要选择同时从事代数和分析学研究的适宜空间. 在通常情形我们取复数域 $\boldsymbol{C}$,它是 $\boldsymbol{Q}$ 对实拓朴完备化域 $\boldsymbol{R}$ 的代数闭包. 而对 p-adic 拓朴,我们类比地取 $\boldsymbol{Q}$ 对 p-adic 拓朴的完备化 $\boldsymbol{Q}_p$ 的代数闭包 $\boldsymbol{Q}_p^{ac}$ 之后,发现 $\boldsymbol{Q}_p^{ac}$ 对 p-adic 拓朴不完备. 于是再取 $\boldsymbol{Q}_p^{ac}$ 的拓朴完备化 C_p 之后,C_p 是代数封闭域,所以 C_p 是我们今后

的活动空间. $\boldsymbol{Q}$ 上的 p-adic 赋值 $|\ |_p$ 和 p-adic 指数赋值 v_p 到 C_p 有惟一扩充,仍表示为 $|\ |_p$ 和 v_p. 例如 $v_p(p)=1$, 从而 $v_p(\sqrt{p})=\frac{1}{2}$, 而 $|\alpha|_p=p^{-v_p(\alpha)}$.

p-adic 赋值具有比三角形不等式 $|a+b|\leqslant|a|+|b|$ 更强的性质 $|a+b|_p\leqslant\max(|a|_p,|b|_p)$,所以有与通常实分析和复分析不同的特性,并且 p-adic 分析常常更加简单. 例如根据定义,C_p 中的柯西序列 $\{a_1,a_2,\cdots,a_n,\cdots\}$ 要求对每个 $\varepsilon>0$,均存在 N,使得当 $n,m\geqslant N$ 时均有 $|a_n-a_m|_p<\varepsilon$. 但实际上,只需要 $|a_{n+1}-a_n|_p\to0$(当 $n\to\infty$ 时)就可以了. 因为当 $m>n$ 时,

$$|a_m-a_n|_p\leqslant\max\{|a_m-a_{m-1}|_p,|a_{m-1}-a_{m-2}|_p,\cdots,$$
$$|a_{n+1}-a_n|_p\}.$$

特别地,C_p 中级数 $\sum\limits_{n=1}^{\infty}a_n$ 收敛当且仅当 $a_n\to0$. 而通常实分析中绝没有这样简单.

再考虑 C_p 上(对 p-adic 拓朴)的连续函数. 它具有任意拓朴域上连续函数的一般性质,比如:连续函数的和、差、积仍是连续函数,由此可知多项式是连续函数. 但是对实拓朴熟知的一些连续函数,考虑 p-adic 拓朴就要仔细. 例如对 $p=3$ 考虑指数函数 $f(x)=2^x$. 取 $x_n=2\cdot3^n$,则 $n\to+\infty$ 时 $x_n\to0$. 而 $2^{x_n}=2^{2\cdot3^n}=2^{\varphi(3^{n+1})}\equiv1\pmod{3^{n+1}}$,即 $f(x_n)\to1$. 另一方面,取 $x'_n=3^n$,则仍有 $x'_n\to0$,但是当 $n\to+\infty$ 时 $2^{x'_n}=2^{3^n}$ 的极限是 $\boldsymbol{Q}_3$ 中的 3 次本原单位根 α,$\alpha\neq1=f(0)$(事实上 $\alpha\equiv2\pmod 3$). 因此 $f(x)=2^x$ 对于 3-adic 拓朴在 $x=0$ 处不连续.

下面是完备距离空间的一个一般性结果(这在拓朴学中是容易证明的),用它来构作 p-adic 连续函数是很实用的.

引理 1 设 S 和 H 是完备距离空间,d_S 和 d_H 为它们的距离函数,A 是 S 的一个稠密子集. 如果函数 $f:A\to H$ 满足如下性质:

对每个 $\varepsilon>0$，均有 $\delta>0$，使得当 $x,y\in A$ 并且 $d_S(x,y)<\delta$ 时就有 $d_H(f(x),f(y))<\varepsilon$，则 f 可以惟一地扩充成从 S 到 H 的一个连续函数. 即存在惟一的连续函数 $F:S\to H$，使得 $F|_A=f$.

证明概要 由于 A 在 S 中稠密，所以对每个 $a\in S$，有 A 中序列 $\{a_n\}$ 使得 $\lim a_n=a$. 由引理中条件可知 $\{f(a_n)\}$ 是 H 中柯西序列，从而 $\lim f(a_n)=\alpha\in H$. 我们定义 $F(a)=\alpha$（这要验证 α 不依赖于 $\{a_n\}$ 的选取方式）. 则 $F:S\to H$ 就是惟一的连续函数，使得 $F|_A=f$.

现在我们用这个引理构作一些 p-adic 连续函数.

例 1 设 n 为正整数，$f(x)=\binom{x}{n}=\frac{1}{n!}x(x-1)\cdots(x-n+1)$，$\binom{x}{0}=1$. 作为多项式，$f(x)$ 是从 $\boldsymbol{Z}_p$ 到 $\boldsymbol{Q}_p$ 的连续函数. 但是当 x 为正整数时，$f(x)\in\boldsymbol{Z}\subseteq\boldsymbol{Z}_p$. 于是 $f:\boldsymbol{Z}\to\boldsymbol{Z}_p$ 是连续函数. 由于 $\boldsymbol{Z}_p$ 是 $\boldsymbol{Q}_p$ 的紧子集，从而 $\boldsymbol{Z}_p$ 对 p-adic 拓朴是完备的，而正整数全体是 $\boldsymbol{Z}_p$ 的稠子集，所以由引理 1 知 $f(x)=\binom{x}{n}$ 实际上是从 $\boldsymbol{Z}_p$ 到 $\boldsymbol{Z}_p$ 的连续函数. 这些连续函数在 p-adic 分析中起着基本作用. 在实分析中，闭区间[0,1]上的连续函数可以用多项式来逼近，从而表成幂级数形式 Σa_nx^n，即 $\{x^n\mid n=0,1,2,\cdots\}$ 形成[0,1]上连续实函数的一组基. 下面的结果表明 $\left\{\binom{x}{n}\middle| n=0,1,\cdots\right\}$ 是紧集 $\boldsymbol{Z}_p$ 上 p-adic 连续函数空间的一组基.

定理 2 (Mahler) 设 K 是 $\boldsymbol{Q}_p$ 的有限扩域，O_K 是 K 的整数环（即 $O_K=\{\alpha\in K\mid v_p(\alpha)\geqslant 0\}$）. 则每个连续函数 $f:\boldsymbol{Z}_p\to O_K$ 都可惟一表示成

$$f(x)=\sum_{n\geqslant 0}a_n\binom{x}{n},$$

其中 $a_n \in O_K, a_n \to 0$.

证明 当 $x \in \mathbf{Z}_p$ 时，$v_p\left(\binom{x}{n}\right) \geqslant 0$，而 $a_n \to 0$，于是 $a_n\binom{x}{n} \to 0$，所以级数 $f(x) = \sum a_n \binom{x}{n}$ 收敛，并且是 $\mathbf{Z}_p$ 上连续函数. 一个初等演算可知对每个 $m \geqslant 0$，

$$a_m = \sum_{i=0}^{m} (-1)^{m-i} \binom{m}{i} f(i).$$

由此可知展开式的惟一性. 证明每个连续函数都可表成定理中形式则比较困难. 从略.

例 2 令 $\alpha \in 1 + p\mathbf{Z}_p$，则 $\alpha = 1 + \beta p (\beta \in \mathbf{Z}_p)$. 对每个非负整数 s，$\alpha^s = (1+\beta p)^s \in 1 + p\mathbf{Z}p$. 于是我们有映射

$$f: \mathbf{N}(\text{非负整数}) \to 1 + p\mathbf{Z}_p, f(s) = \alpha^s.$$

如果 $s, s' \in \mathbf{N}$ 并且 $v_p(s - s') \geqslant M$，则 $s' = s + p^M s'' (s'' \in \mathbf{Z})$. 不妨假设 $s' \geqslant s$（即 $s'' \geqslant 0$），则

$$v_p(\alpha^s - \alpha^{s'}) = v_p(\alpha^s) + v_p(1 - \alpha^{p^M s''}) \geqslant v_p(1 - \alpha^{p^M s''}).$$

但是

$$\alpha^{p^M s''} = (1 + \beta p)^{p^M s''} \equiv 1 \pmod{p^{M+1}},$$

所以 $v_p(\alpha^s - \alpha^{s'}) \geqslant M + 1$. 这就表明 f 满足引理 1 中条件（$\mathbf{N}$ 为 $\mathbf{Z}_p$ 的稠密子集）. 于是由 f 可定义出 p-adic 连续函数（其中 $\alpha \in 1 + p\mathbf{Z}_p$）

$$f: \mathbf{Z}_p \to 1 + p\mathbf{Z}_p, \quad f(s) = \alpha^s.$$

由上面推导可知，对于 $s = a_0 + a_1 p + \cdots + a_n p^n + \cdots \in \mathbf{Z}_p (0 \leqslant a_i \leqslant p - 1)$，取正整数 $s' = a_0 + \cdots + a_N p^N$，则 $\alpha^{s'}$（即 s' 个 α 相乘）为 α^s 的近似值，并且精确到 p-adic 展开的第 p^{N+1}-位、性质 $\alpha^{s+s'} = \alpha^s \cdot \alpha^{s'}$ 和 $\alpha^{-s} = (\alpha^s)^{-1}$ 在 $s, s' \in \mathbf{N}$ 时成立. 由函数的连续性知在 $\mathbf{Z}_p$ 中也成立.

更一般地，如果 $\alpha \in U_p = \{\alpha \in \mathbf{Z}_p \mid v_p(\alpha) = 0\}$（p-adic 单位群），

则 $\alpha^{p-1}\in 1+p\boldsymbol{Z}_p$. 对每个 $a=0,1,2,\cdots,p-2$,考虑集合

$$A_a=a+(p-1)\boldsymbol{N}=\{a+(p-1)n\,|\,n=0,1,2,\cdots\},$$

易知每个 A_a 仍是 $\boldsymbol{Z}_p$ 的稠密子集. 对每个 a 考虑函数

$$f_a:A_a\to\boldsymbol{Z}_p, f_a(a+(p-1)n)=\alpha^{a+(p-1)n}=\alpha^a\cdot(\alpha^{p-1})^n.$$

如果 $v_p((a+(p-1)n)-(a+(p-1)n'))=v_p((p-1)(n-n'))\geqslant M$, 则 $v_p(n-n')\geqslant M$. 由 $\alpha^{p-1}\in 1+p\boldsymbol{Z}_p$ 可知 $v_p(f_a(a+(p-1)n)-f_a(a+(p-1)n'))\geqslant M+1$. 于是由引理 1,对每个 $\alpha\in U_p$,我们构作了 $p-1$ 个不同的 p-adic 连续函数

$$f_a:\boldsymbol{Z}_p\to\boldsymbol{Z}_p\quad(a=0,1,2,\cdots,p-2).$$

f_a 在 A_a 上是通常意义下的指数函数,即对于 $s\in A_a$, $f_a(s)=\alpha^s$ 为 s 个 α 相乘. 当 $1\leqslant a\leqslant p-2$ 时,指数函数的性质 $f_a(s+s')=f_a(s)f_a(s')$ 不必成立,因在 $s,s'\in A_a$ 时, $s+s'\notin A_a$. 但是 $f_0(s+s')=f_0(s)f_0(s')$ 在 A_0 上成立. 由连续性知在 $\boldsymbol{Z}_p$ 上也成立.

$\boldsymbol{Z}_p$ 上每个 p-adic 连续函数都可惟一展成 $\Sigma a_n\binom{x}{n}$ (Mahler 定理),但是也像实分析中那样,常常用幂级数来定义连续函数. 设

$$f(x)=\sum_{n\geqslant 0}a_nx^n\in C_p[[x]]$$

是系数属于 C_p 的幂级数. 对每个 $\alpha\in C_p$, Σa_nx^n 收敛当且仅当 $|a_nx^n|_p\to 0$. 和实分析一样,令

$$r=\left(\varlimsup_{n\to\infty}|a_n|_p^{\frac{1}{n}}\right)^{-1},$$

则当 $|\alpha|_p<r$ 时, $\Sigma a_n\alpha^n$ 收敛,而当 $|\alpha|_p>r$ 时此级数发散,并且 $f(x)=\Sigma a_nx^n$ 是开球

$$D(r^-)=\{x\in C_p\,|\,|\alpha|_p<r\}$$

中的连续函数. r 叫此幂级数的**收敛半径**. 进而,对于 $f'(x)=\sum_{n\geqslant 1}na_nx^{n-1}$. 由于 $|na_n|_p\leqslant|a_n|_p$, 可知 $f'(x)$(从而 $f(x)$ 的任何阶微商)在 $D(r^-)$ 中仍是连续函数,即 $f(x)$ 在 $D(r^-)$ 中是无限次可

微的解析函数.

例 3 p-adic 指数函数

$$\exp(x)=\sum_{n\geqslant 0}\frac{x^n}{n!}$$

对于复分析情形，$r=\lim\limits_{n\to\infty}(n!)^{\frac{1}{n}}=+\infty$，所以 $\exp(x)$ 是 $\boldsymbol{C}$ 上的解析函数. 但是对 p-adic 情形，$n!$ 中 p 的指数为 $\left[\frac{n}{p}\right]+\left[\frac{n}{p^2}\right]+\cdots$，所以当 n 很大时，$\left|\frac{1}{n!}\right|_p$ 可以很大. 由于 $v_p(n!)=\left[\frac{n}{p}\right]+\left[\frac{n}{p^2}\right]+\cdots<\frac{n}{p-1}$，而当 $p^a\leqslant n<p^{a+1}$ 时，

$$\begin{aligned}v_p(n!)&>\frac{n}{p}+\frac{n}{p^2}+\cdots+\frac{n}{p^a}-a\\&=\frac{n}{p-1}-a-\frac{np^{-a}}{p-1}>\frac{n-p}{p-1}-\frac{\log n}{\log p},\end{aligned}$$

可知 $\lim\frac{1}{n}v_p(n!)=\frac{1}{p-1}$，即 $\lim|n!\ |_p^{\frac{1}{n}}=p^{-\frac{1}{p-1}}$. 即 $\exp(x)$ 的收敛半径为 $p^{-\frac{1}{p-1}}$.

例 4 p-adic 对数函数

$$\log_p(1+x)=\sum_{n\geqslant 1}\frac{(-1)^{n+1}}{n}x^n$$

在实分析中它的收敛半径为 1，因为 $\lim n^{\frac{1}{n}}=1$. 对于 p-adic 情形，由于 n 至多被 $\frac{\log n}{\log p}$ 个 p 除尽，所以 $0\leqslant v_p(n)\leqslant\frac{\log n}{\log p}$，$\lim\frac{1}{n}v_p(n)=0$，即 $\log_p(1+x)$ 的收敛半径也为 1.

令 $q=p$（若 p 为奇素数），$q=4$（若 $p=2$）. 由例 3 和例 4 知 $\exp(x)$ 是 $q\boldsymbol{Z}_p$ 中解析函数，并且 $\exp(q\boldsymbol{Z}_p)\subseteq 1+p\boldsymbol{Z}_p$，而 $\log_p(1+x)$ 是 $p\boldsymbol{Z}_p$ 中解析函数，并且 $\log_p(1+p\boldsymbol{Z}_p)\subseteq q\boldsymbol{Z}_p$. 进而有

$$\exp\circ\log_p(1+x)=1+x\quad(x\in p\boldsymbol{Z}_p),$$
$$\log_p\circ\exp(x)=x\quad(x\in q\boldsymbol{Z}_p).$$

从而我们有互逆的映射

$$qZ_p \underset{\log_p}{\overset{\exp}{\rightleftarrows}} 1+pZ_p.$$

还可证明

$$\exp(x+y)=\exp(x)\cdot\exp(y) \quad (x,y\in qZ_p),$$
$$\log_p[(1+x)(1+y)]$$
$$=\log_p(1+x)+\log_p(1+y)(x,y\in pZ_p).$$

所以给出加法群 qZ_p 和乘法群 $1+pZ_p$ 的同构. 以上诸公式都可按这样方式证明:它们在实分析中是对的,所以作为形式幂级数也是对的,于是在 p-adic 收敛区域内也是对的.

乘法群 Q_p^x 有如下的直和分解

$$Q_p^x=\langle p\rangle\times W\times(1+pZ_p), U_p=W\times(1\oplus pZ_p),$$

其中 $\langle p\rangle$ 是由 p 生成的无限循环群, W 是 Z_p 中的 $p-1$ 次单位根形成的群,以 ζ 表示满足 $\zeta\equiv1(\mathrm{mod}\ \mathrm{p})$ 的 $\mathrm{p}-1$ 次本原单位根,则 Q_p^x 中元素 α 惟一表成

$$\alpha=p^n\zeta^i\langle\alpha\rangle(n=v_p(\alpha)\in Z, 0\leqslant i\leqslant p-2, \langle\alpha\rangle\in1+pZ_p).$$

我们定义 $\log_p\alpha=\log_p\langle\alpha\rangle$(即规定 $\log_p p=\log_p\zeta=1$),这就把 $\log_p$ 定义到整个 Q_p^x 上,并且仍满足 $\log_p(\alpha\beta)=\log_p\alpha+\log_p\beta$.

现在对 $\alpha\in U_p$,可以定义函数

$$\langle\alpha\rangle^x=\exp(x\log_p\langle\alpha\rangle)=\exp(x\log_p\alpha).$$

由于 $\log_p\langle\alpha\rangle\in qZ_p$, 所以 $\langle\alpha\rangle^x$ 是 $\{x\in C_p \mid |x|_p<qp^{-\frac{1}{p-1}}\}$ 中的连续函数. 当 $n\in Z$ 时 $\langle\alpha\rangle^n$ 就是通常的意义, 而在 $n\equiv0(\mathrm{mod}\ p-1)$ $(p\geqslant3)$ 和 $n\equiv0(\mathrm{mod}\ 2)(p=2)$ 时, $\langle\alpha\rangle^n=\alpha^n$.

例 5 (p-adic zeta 函数)19 世纪人们对于贝努利数 B_n(这是有理数)的整性和同余性得出如下结果

引理 3(von Staudt-Clausen) 对每个正偶数 n 和素数 p,当 $p-1\nmid n$ 时 $B_n\in Z_p$,而当 $p-1\mid n$ 时, $B_n+\frac{1}{p}\in Z_p$.

引理 4(Kummer)　设 $N\geqslant 1, 2k\not\equiv 0 \pmod{p-1}, 2k\equiv 2k' \pmod{(p-1)p^N}$. 则

$$-(1-p^{2k-1})\frac{B_{2k}}{2k}\equiv -(1-p^{2k'-1})\frac{B_{2k'}}{2k'} \pmod{p^{N+1}}.$$

并且同余式两边均是 p-adic 整数.

一百年后，人们才以新的观点看待库默尔这个同余式：这不恰好可以用来定义 p-adic 连续函数吗？设 $p\geqslant 5$，对每个正整数 a，$1\leqslant a\leqslant \frac{p-3}{2}$，定义集合

$$\begin{aligned} C_a &= \{1-(2a+(p-1)b) \mid b=0,1,2,\cdots\} \\ &= 1-2a-(p-1)\mathbf{N}. \end{aligned}$$

每个 C_a 都是 $\mathbf{Z}_p$ 的稠密子集，并且 $\frac{p-3}{2}$ 个集合 C_a 彼此不相交. 考虑映射

$$f_a: C_a \to \mathbf{Z}_p, f_a(1-2k) = -(1-p^{2k-1})\frac{B_{2k}}{2k}.$$

引理 4 恰好是说：若 $1-2k, 1-2k'\in C_a$，$v_p((1-2k)-(1-2k'))=v_p(2k'-2k)\geqslant N$，则 $v_p(f_a(1-2k)-f_a(1-2k'))\geqslant N+1$. 所以由引理 1，它可扩充为 p-adic 连续函数 $f_a:\mathbf{Z}_p\to\mathbf{Z}_p$.

我们知道，黎曼 zeta 函数 $\zeta(n)=\sum\limits_{n\geqslant 1} n^{-s}$ 解析延拓到整个复平面上之后，它在负奇数处的值为 $\zeta(1-2k)=-\frac{B_{2k}}{2k}$. 如果考虑级数 $\sum n^{-s}$ 的 p-adic 收敛性，它显然是发散的，因为对 $n=p^l$，$|n^{-s}|_p=p^{sl}$ 很大(当 $s>0$ 时). 我们需要把级数中满足 $p|n$ 的那些项 n^{-s} 去掉，即令

$$\begin{aligned} \zeta^*(s) &= \sum_{\substack{n\geqslant 1 \\ p\nmid n}} n^{-s} = \prod_{\substack{q \\ q\neq p}} (1-q^{-s})^{-1} \\ &= (1-p^{-s})\zeta(s) \quad (s>1), \end{aligned}$$

其中 q 过 p 以外的所有素数. 这时

$$\zeta^*(1-2k)=(1-p^{2k-1})\zeta(1-2k)$$

$$=-(1-p^{2k-1})\frac{B_{2k}}{2k}\quad (k=1,2,\cdots).$$

所以 p-adic 连续函数 f_a 在 C_a 上取值和 ζ^* 一致. 我们把 $f_a(s)$ 表示成 $\zeta_p^{(a)}(s)$, 叫作 p-adic zeta 函数. 这种看法和构作方式是 1964 年由 Leopoldt(Hasse 的学生)和久保田(Kubota)给出的. 于是对每个 $p\geqslant 5$, 我们有 $\frac{p-3}{2}$ 个 p-adic 连续函数(从 $\mathbf{Z}_p$ 到 $\mathbf{Z}_p$) $\zeta_p^{(a)}(s)$ $\left(1\leqslant a\leqslant\frac{p-3}{2}\right)$, 其中 $\zeta_p^{(a)}(s)$ 和 $\zeta^*(s)$ 在 $\mathbf{Z}_p$ 的稠密子集合 C_a 上取值一样.

类似地也可构作 p-adic$L$-函数, 其结果为

定理 5 设 p 为素数, χ 是导子为 f 的本原狄里赫利特征. $q=p$(当 $p\geqslant 3$), $q=4$(当 $p=2$ 时). 对每个正整数 F, $q|F$, $f|F$, 定义

$$L_p(s,\chi)$$
$$=\frac{1}{F(s-1)}\sum_{\substack{a=1\\ p\nmid a}}^{F}\chi(a)\langle a\rangle^{1-s}\sum_{j\geqslant 0}\binom{1-s}{j}\left(\frac{F}{a}\right)^j B_j\in C_p.\quad (*)$$

若 χ 不是主特征, 则 $L_p(s,\chi)$ 是 $D=\{s\in C_p\mid |s|_p\leqslant qp^{-\frac{1}{p-1}}\}$ 中的解析函数. 若 χ 为主特征, 则 $s=1$ 为 $L_p(s,\chi)$ 的单极点并且留数为 $1-\frac{1}{p}\left(\text{即}\lim\limits_{s\to 1}(s-1)L_p(s,\chi)=1-\frac{1}{p}\right)$, 而在 $1\neq s\in D$ 时均解析. 并且, $L_p(s,\chi)$ 和函数

$$L^*(s,\chi)=\sum_{\substack{n\geqslant 1\\ p\nmid n}}\chi(n)n^{-s}=(1-\chi(p)p^{-s})L(s,\chi)$$

在集合 $\{1-2k\mid k=1,2,3,\cdots\}$ 上取值相同, $L_p(s,\chi)$ 叫作 p-adic L-函数.

公式($*$)中的 $\langle a\rangle$ 由本节讲述 $\log_p$ 时所定义, $\chi(a)$(0 或 f 次单位根)看作 C_p 中元素.

这些 p-adic L-函数有什么用途呢？对于古典情形我们有狄里赫利 L-函数在 $s=1$ 值 $L(1,\chi)$的计算公式，并且阿贝尔数域 K 的类数 $h(k)$可以表示成一些 $L(1,\chi)$的乘积. 对于 p-adic 情形我们有类似的公式.

定理 6 设 K 为 n 次实阿贝尔域，$\hat{K}$ 是伽罗华群 Gal(K/Q)的特征群. 则

$$\frac{2^{n-1}h(K)R_p(K)}{\sqrt{d(K)}}=\prod_{1\neq\chi\in\hat{K}}\left(1-\frac{\chi^*(p)}{p}\right)^{-1}L_p(1,\chi^*)\in C_p,$$

其中 χ^* 表示与 χ 对应的本原特征，$R_p(K)$与 $R(K)$的定义相仿但属于 C_p，叫域 K 的 p-adic regulator. 进而，也有与古典情形相仿的 $L_p(1,\chi)$公式(普通对数 log 改成 p-adic 对数 $\log_p$).

上面公式叫作 p-adic 类数公式. 通常类数公式中出现超越数 $R(K)$和 $\log(1-\zeta^a)$，在复数域中对复拓朴计算类数 $h(K)$和它的近似值比较困难. 而用p-adic类数公式，$R(K)$和 $\log_p$ 的 p-adic 近似值容易得到，从而容易计算类数 $h(K)$的 p-adic 近似值. 所以 p-adic 方法特别适合于研究类数的整除性质. 比如说，库默尔关于分圆域类数的整除性结果(见 § 2.2 库默尔结果(5)和(6))以及在 p-adic L-函数出现之前人们用有限群表示论得到的关于分圆域和阿贝尔域一系列重要结果(见下节)都可用 p-adic L-函数和 p-adic 分析方法统一地证明，并且对这些结果给予新的看法和更深的理解. 当然，一个方法的价值主要是看能否得到新的重要结果和猜想. 用 p-adic L-函数得到的第一个重要结果是由岩泽健吉(Iwasawa)给出的，其结果涉及由无限个数域组成的扩张序列中类数的变化性状.

一个数域的扩张序列

$$K=K_0\subset K_1\subset K_2\subset\cdots K_n\subset\cdots$$

如果对每个 $n\geqslant 0$，K_n/K 是 p^n 次循环扩张，我们称这个扩张序列

为 K 的 $\mathbf{Z}_p$-扩张. 这是因为令 $K_\infty=\bigcup_{n\geqslant 0}K_n$，则 K_∞/K 是无限伽罗华扩张，它的伽罗华群 $\mathrm{Gal}(K_\infty/K)$ 是 $\mathrm{Gal}(K_n/K)\cong\mathbf{Z}/p^n\mathbf{Z}$ 的逆极限，所以同构于加法群 $\varprojlim_n \mathbf{Z}/p^n\mathbf{Z}=\mathbf{Z}_p$. 反过来. 若 K_∞/K 是无限伽罗华扩张，并且 $\mathrm{Gal}(K_\infty/K)$ 同构于加法群 $\mathbf{Z}_p$. 根据无限伽罗华扩张的基本定理，K_∞/K 的中间域和 $\mathbf{Z}_p$ 的闭子群一一对应. 由于 $\mathbf{Z}_p$ 的闭子群只有 $p^n\mathbf{Z}_p(n=0,1,2,\cdots)$，所以 K_∞/K 的所有中间域构成上述序列，并且 $\mathrm{Gal}(K_n/K)$ 同构于 $\mathbf{Z}_p/p^n\mathbf{Z}_p\cong\mathbf{Z}/p^n\mathbf{Z}$，即 K_n/K 是 p^n 次循环扩张，所以我们也把 K_∞/K 叫作 $\mathbf{Z}_p$-扩张. 例如对每个奇素数 p，令 $K_n=\mathbf{Q}(\zeta_{p^{n+1}})(n=0,1,2,\cdots)$，则

$$\begin{aligned}\mathrm{Gal}(K_n/K_0)&\cong\mathrm{Gal}(K_n/\mathbf{Q})/\mathrm{Gal}(K_0/\mathbf{Q})\\&\cong(\mathbf{Z}/p^{n+1}\mathbf{Z})^\times/(p-1\text{ 阶子群})\cong\mathbf{Z}/p^n\mathbf{Z}.\end{aligned}$$

所以 $K_0=\mathbf{Q}(\zeta_p)\subset K_1\subset\cdots\subset K_n\subset\cdots$ 是 $\mathbf{Q}(\zeta_p)$ 的 $\mathbf{Z}_p$-扩张. 如果以 L_n 表示 K_n 的惟一子域使得 $[K_n:L_n]=p-1$，则

$$L_0=\mathbf{Q}\subset L_1\subset\cdots\subset L_n\subset\cdots$$

便是 $\mathbf{Q}$ 的 $\mathbf{Z}_p$-扩张. 类似地，$K_n=\mathbf{Q}(\zeta_{2^{n+2}}+\bar{\zeta}_{2^{n+2}})$ 是 $\mathbf{Q}$ 的 2^n 次循环扩张，从而 $K_0=\mathbf{Q}\subset K_1\subset\cdots\subset K_n\subset\cdots$ 是 $\mathbf{Q}$ 的 $\mathbf{Z}_2$-扩张. 而令 $L_n=\mathbf{Q}(\zeta_{2^{n+2}})$，则 $L_0\subset L_1\subset\cdots\subset L_n\subset\cdots$ 是 $L_0=\mathbf{Q}(i)$ 的 $\mathbf{Z}_2$-扩张. 更一般地，对任意数域 K，令 $K_n=K(\zeta_{p^n})$，则存在 $l\geqslant 0$，使得 $\zeta_{p^l}\in K$，$\zeta_{p^{l+1}}\notin K$. 于是 $K_l\subset K_{l+1}\subset\cdots\subset K_{l+n}\subset\cdots$ 是 K_l 的 $\mathbf{Z}_p$-扩张. 这类 $\mathbf{Z}_p$-扩张都叫作分圆 $\mathbf{Z}_p$-扩张.

定理 7(岩泽) 设 K 为数域，$K=K_0\subset K_1\subset\cdots\subset K_n\subset\cdots$ 是 K 的 $\mathbf{Z}_p$-扩张. 对素数 p，以 e_n 表示最大非负整数使得 $p^{e_n}\mid h(K_n)$(K_n 的理想类数)，即 $e_n=v_p(h(K_n))$. 则存在和 n 无关的非负整数 μ,λ 和整数 ν，使得对每个 $n\geqslant 0$ 均有

$$e_n=\mu p^n+\lambda n+\nu.$$

这是一个深刻的结果. 它依赖于更多的代数工具，也依赖于岩

泽给出构作 p-adic L-函数的新方法:幂级数方法.他把每个 p-adic L-函数对应于系数属于 $\mathbf{Z}_p$ 的关于 T 的幂级数,即幂级数环 $\Lambda=\mathbf{Z}_p[[T]]$中的元素.

岩泽对环 Λ 的研究来源于数域的 $\mathbf{Z}_p$-扩张:$k=k_0\subset k_1\subset\cdots\subset k_n\subset\cdots$.令 $\Gamma=\mathrm{Gal}(k_\infty/k)$,它同构于加法群 $\mathbf{Z}_p$.作为拓朴群 $\mathbf{Z}_p$ 是由 1 生成的无限循环群 $\mathbf{Z}$ 的拓朴闭包.因此 Γ(乘法群)也是由某元素 γ 生成的无限循环群的拓朴闭包.K_n 的固定子群为 Γ^{p^n}(由 γ^{p^n} 生成的无限循环群的拓朴闭包),于是 $\Gamma_n=\mathrm{Gal}(K_n/K)=\Gamma/\Gamma^{p^n}$是 p^n 阶循环群.当 $m\geqslant n\geqslant 0$ 时由自然映射 $\Gamma_m\to\Gamma_n$ 诱导出环同态 $\mathbf{Z}_p[\Gamma_m]\to\mathbf{Z}_p[\Gamma_n]$.而显然 $\mathbf{Z}_p[\Gamma_n]\cong\mathbf{Z}_p[T]/((1+T)^{p^n}-1)$(由 $\gamma\,(\mathrm{mod}\,\Gamma^{p^n})\mapsto 1+T(\mathrm{mod}((1+T)^{p^n}-1))$,于是给出

$$\mathbf{Z}_p[\Gamma]=\varprojlim\mathbf{Z}_p[\Gamma_n]\cong\varprojlim\mathbf{Z}_p[T]/((1+T)^{p^n}-1).$$

而后者同构于 $\Lambda=\mathbf{Z}_p[[T]]$.最后得到

$$\mathbf{Z}_p[\Gamma]\simeq\Lambda=\mathbf{Z}_p[[T]],\text{其中 }\gamma\mapsto 1+T.$$

$\mathbf{Z}_p[[T]]$不是主理想整环,但它是惟一因子分解整环.元素 $f(T)=\sum_{n=0}^{\infty}a_nT^n\in\Lambda$ 为单位当且仅当 a_0 是 $\mathbf{Z}_p$ 中单位(即 $v_p(a_0)=0$).而不可约元素可取为 p 和一类特殊的多项式 $f(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$,其中 $v_p(a_i)\geqslant 1(0\leqslant i\leqslant n-1)$,叫作 distinguished 多项式.换句话说,环 Λ 有如下的分解特性

引理 8 (p-adic 外尔斯特拉斯分解定理) 对每个 $0\neq f(T)\in\Lambda=\mathbf{Z}_p[[T]]$,则 $f(T)$可惟一分解成

$$f(T)=p^\mu P(T)U(T),$$

其中 μ 为非负整数,$U(T)$为 Λ 中单位,$P(T)$为 distinguished 多项式.

利用 Λ 的这个分解定理,岩泽证明了有限生成 Λ-模的结构

定理.

定理 9(岩泽) 每个有限生成Λ-模均拟同构于$\Lambda^r \oplus \bigoplus_{i=1}^{s} \Lambda/(p^{n_i}) \oplus \bigoplus_{j=1}^{t} \Lambda/(f_j(T))^{m_j}$,其中$r,t,s,n_i,m_j$均为非负整数,$f_j(T)$为$\Lambda$中不可约的 distinguished 多项式.而Λ-模M和N叫作**拟同构**,是指存在Λ-模同态$\varphi: M\to N$,使得$\ker\varphi$和$\operatorname{coker}\varphi=N/I_m\varphi$都是有限的.

现在介绍定理 7 的证明思想:对每个$n\geqslant 0$,记L_n为K_n的最大不分歧阿贝尔p-扩张.根据类域论,$X_n=\mathrm{Gal}(L_n/K_n)$同构于$k_n$的理想类群的 Sylow p-子群A_n.再令L_∞是$K_\infty=\bigcup_{n\geqslant 0} K_n$的最大不分歧阿贝尔$p$-扩张,则$L_\infty=\bigcup_{n\geqslant 0} L_n$.$\Gamma=\mathrm{Gal}(K_\infty/K)$自然作用在$X=\mathrm{Gal}(L_\infty/K_\infty)$上,由于$X$是射影$p$-群,从而$\mathbf{Z}_p$也自然作用于$X$上.于是$X$为$\mathbf{Z}_p[\Gamma]$-模.而$\mathbf{Z}_p[\Gamma]\cong\mathbf{Z}_p[[T]]=\Lambda$,因此$X$是$\Lambda$-模.进一步,岩泽证明了$X$是 torsion Λ-模.由结构定理 9,可知X拟同构于$\bigoplus_{i=1}^{s}\Lambda/(p^{n_i})$和$\bigoplus_{j=1}^{r}\Lambda/(f_j(T))$的直和,其中$f_j(T)$是 distinguished 多项式.然后再将$X$投射到$X_n$,便可算出$A_n(\cong X_n)$的阶$p^{e_n}$,从而得到定理 7 中关于$e_n$的公式.

上述证明中并没有出现 p-adic L-函数.但是岩泽首先证明定理 7 的一个最简单情形时采用了 p-adic L-函数方法,即利用分圆域$\mathbf{Q}(\zeta_{p^n})$类数h_{p^n}的 p-adic 解析公式,把相对类数$h_{p^n}^-$表示成一些$L_p(1,\chi)$的乘积,再用Λ中的幂级数对应于每个 p-adic L-函数$L_p(s,\chi)$,然后利用Λ中幂级数的分解特性(引理 8)对于$h_{p^n}^-$给出了类似于定理 7 的结果.

岩泽关于数域$\mathbf{Z}_p$-扩张的类数公式$e_n=\mu p^n+\lambda n+\nu$是受到函数域情形的启发.设$C$是$\mathbf{F}_q$上绝对不可约的非奇异曲线.根据§3.4,我们对函数域$K=\mathbf{F}_q(C)$的除子类数$h(K)$有公式$h(K)=$

$L(1)$，其中 $L(U)=\prod_{i=1}^{2g}(1-w_iU)\in\mathbf{Z}[U]$ 是 zeta 函数 $Z_k(U)=L(U)/(1-U)(1-qU)$ 的分子多项式，g 为曲线 C 的亏格．令 $K_n=\boldsymbol{F}_{q^n}(C)$（$K$ 的系数域扩张），则 K_n/K 为伽罗华扩张，其伽罗华群同构于 $\mathrm{Gal}(\boldsymbol{F}_{q^n}/\boldsymbol{F}_q)$，这是 n 次循环扩张．利用 k 的素理想在 k_n 中的分解规律，可得到 K_n 的 zeta 函数为 $Z_{K_n}(U)=L_n(U)/(1-U)(1-q^nU)$，其中 $L_n(U)=\prod_{i=1}^{2g}(1-w_i^nU)\in\mathbf{Z}[U]$，而 $h(K_n)=L_n(1)$．现在设 p 是 q 的素因子，取 $K'_n=K_{p^n}$，则 $K=K'_0\subset K'_1\subset\cdots\subset K'_n\subset\cdots$ 为 $\mathbf{Z}_p$-扩张．令 $e_n=v_p(h(K'_n))$，利用上面的关系可证当 n 充分大时有 $e_n=\lambda n+\nu$，其中 λ,ν 与 n 无关．关于数域 $\mathbf{Z}_p$-扩张情形比函数域复杂，需要 Λ-模的深刻结果，并且公式中多出一项 μp^n．但是基于这种数域和函数域的类比，岩泽提出如下的

猜想　对每个数域 K 和素数 p，关于 K 的分圆 $\mathbf{Z}_p$-扩张的（岩泽）不变量 μ 均为 0.

1979 年，Ferrero 和 L. Washington 对于 K 是阿贝尔域的情形证明了这个猜想．Greenberg 进一步猜想：若 K 是全实域，则对 K 的分圆 $\mathbf{Z}_p$-扩展，岩泽不变量 λ 也为 0. 这两个猜想至今都未解决．

在研究分圆域和 $\mathbf{Z}_p$-扩张过程中产生的构作 p-adic L-函数方法和 Λ-模理论，对近代数论其他分支产生很大影响．为了研究模形式 Fourier 展开系数的同余和整除性质，发展了 p-adic 模形式理论．对于定义在数域 k 上椭圆曲线 E 和 K 的 $\mathbf{Z}_p$-扩张 $\{k_n\}$ $(n=0,1,2,\cdots)$，为了研究点群 $E(K_n)$ 的性状，人们构作了椭圆曲线 E 的 p-adic L-函数 $L_p(E,S)$ 和相关联的 Λ-模结构．我们在上节介绍的 Coates-wiles 和 K. Rubin 关于椭圆曲线 BSD 猜想的重要结果就采用了 p-adic L-函数 $L_p(E,s)$ 作为重要工具之一．反过来，模形式和椭圆曲线的这种 p-adic 理论又促进分圆域理论的发展．

Klingen 和 Siegel 分别于 1962 年和 1969 年证明了：对每个全实域 K（即 K 到 $\boldsymbol{C}$ 的嵌入均是实嵌入），狄德金 zeta 函数 $\zeta_K(s)$在非负整数的取值均为有理数，方法是它们是某些希尔伯特模形式 Fourier 展开的常数项，就像黎曼 zeta 函数值是爱森斯坦级数的常数项一样. 然后 Deligne 和 Ribet 于 1980 年利用 p-adic 希尔伯特模形式的结果得到这些有理数的同余性质，从而作为插值给出全实域 k 的 p-adic L-函数 $L_p(s,\chi)$，它在一批负整数上取值和 $L(s,\chi)=\sum_A \chi(A)NA^{-s}$ 一致，其中 A 过 O_K 的整理想，而 χ 是实阿贝尔扩张 M/K 的伽罗华群的特征（根据类域论，χ 可看成是 K 的某个广义类群的特征）. 设 $F=K(\zeta_p)$，F_∞是 F 的分圆 $\boldsymbol{Z}_p$-扩张（即 $F_n=K(\zeta_{p^n})$）. 则 $\Delta=\mathrm{Gal}(F/K)$（$p-1$阶循环群同构于$(\mathbf{Z}/p\mathbf{Z})^x$）可看成模 p 狄里赫利特征. 设 w 为模 p 的 Teichmüller 特征，即对与 p 互素的整数 a，$w(a)$为满足 $w(a)\equiv a(\mathrm{mod}\ p)$的 $p-1$ 次本原单位根，对于 Δ 的每个奇特征 χ，Barsky(1977)，Cassou-Noguès 和 Deligne-Ribet(1980)都证明了存在 Λ 中形式幂级数 f_χ 对应于 F 的 p-adic L-函数 $L_p(s,w\chi^{-1})$，即 $L_p(s,w\chi^{-1})=f_\chi((1+p)^s-1)$. 问题在于如何得到 $f_\chi(T)$？在这方面，岩泽有如下的猜想.

以 L_∞表示 F_∞的最大不分歧阿贝尔 p-扩张，如前所述，$X=\mathrm{Gal}(L_\infty/F_\infty)$为 torsion Λ-模，对于群环 $\boldsymbol{Z}_p[\Delta]$的幂等元 $\varepsilon_\chi=\frac{1}{|\Delta|}\sum_{\delta\in\Delta}\chi^{-1}(\delta)\delta$，$\varepsilon_\chi X$ 也是 torsion Λ-模. 于是拟同构于$\bigoplus_i \Lambda/(p^{n_i})$和$\bigoplus_j \Lambda/(g_j(T))$的直和，令 $\mu=\sum_i n_i$，$g(T)=p^\mu\prod_j g_j(T)$.

岩泽主猜想 设 χ 为 Δ 的奇特征，$\chi\neq w$，则 $f_\chi(T)=g(T)\cdot U(T)$，其中$U(T)$是 Λ 中单位.

这是近代分圆域理论一个重要猜想. 第一个突破是由 Mazur 和 Wiles 于 1984 年得到的，他们对 $K=\boldsymbol{Q}$，$F=\boldsymbol{Q}(\zeta_p)$的情形证明

了岩泽主猜想，采用了与某些模形式相联系的 p-adic 表示. 后来 K. Rubin 于 1990 年用 Kolyvagin 在椭圆曲线理论中创造的 Euler system 方法给出一个简化的证明，见 L. C. Washington《Introduction to Cyclotomic Fields》一书第二版的第十五章，这个结果在分圆域理论和代数 K-理论中都有重要的推论. 1992 年，Greither 采用 Kolyvagin 和 Rubin 方法对 K 为任意阿贝尔域的情形证明了岩泽主猜想.

以上我们介绍了构作 p-adic zeta 函数和 p-adic L-函数的插值方法和幂级数方法. 现在介绍由 Mazur 于 1972 年创造，后来被广泛采用的第三种方法，即 p-adic 测度和 p-adic 积分方法. 为简单起见，我们以 $\boldsymbol{Q}_p$ 上的积分为例. 但是首先要回忆一下实数域 $\boldsymbol{R}$ 上的积分是如何定义的.

设 μ 是区间[0,1]上的一个测度. 这意味着：对[0,1]的每个可测子集合 M，$\mu(M)$为非负实数，并且满足如下的可加性：若 M_i $(i\in I)$为[0,1]的任意多个可测子集合，并且彼此不相交，则

$$\mu(\bigcup_{i\in I} M_i) = \sum_{i\in I}\mu(M_i).$$

这时，对[0,1]上任意连续函数 $f(x)$，我们把[0,1]分成一些小区间

$$I_1=[x_0,x_1),I_2=[x_1,x_2),\cdots,I_n=[x_{n-1},x_n],$$

其中 $0=x_0<x_1<\cdots<x_n=1$. 任取 $x'_i\in[x_{i-1},x_i](1\leqslant i\leqslant n)$，考虑

$$\sum_{i=1}^{n} f(x'_i)\mu(I_i).$$

这个求和在区间分得愈来愈细时趋于极限，这个极限叫作 $f(x)$关于测度 μ 在[0,1]上的积分，表示成 $\int_0^1 f(x)\mu(x)$.

现在考虑 $\boldsymbol{Q}_p$. 我们把紧开子集 $\boldsymbol{Z}_p$ 类比于 $\boldsymbol{R}$ 的闭区间[0,1]. 更一般的，对每个 $a\in\boldsymbol{Q}_p,N\in\boldsymbol{Z}$，我们把子集合

$$a+(p^N)=a+p^N\boldsymbol{Z}_p=\{x\in\boldsymbol{Q}_p\,|\,v_p(x-a)\geqslant N\}$$

叫作是 $\boldsymbol{Q}_p$ 的一个**区间**，N 叫此区间的**级别**. 例如：$p^N\boldsymbol{Z}_p=0+$

(p^N). 下面引理表明它们在 p-adic 拓朴中起着基本作用.

引理 10 $\boldsymbol{Q}_p$ 的每个紧开子集都是彼此不相交的有限个同级别区间的并集.

证明大意 每个紧开子集 U 都是有限个区间的并集. 再由级别为 N 的区间 $a+(p^N)$ 可分成 p 个级别为 $N+1$ 的区间 $a+jp^N+(p^{N+1})(0\leqslant j\leqslant p-1)$ 之并, 所以 U 一定是有限个同级别的区间之并. 但是同级别的两个区间或者不相交, 或者完全相等(这是 p-adic 拓朴的特征). 所以把重合的区间去掉之后, U 就是彼此不相交的有限个同级别区间之并集.

设 X 是 $\boldsymbol{Q}_p$ 的紧开子集, μ 是一个映射, 把 X 的每个紧开子集 U 映成 $\boldsymbol{Q}_p$ 中元素 $\mu(U)$. 我们称 μ 是 X 上的一个 p-adic **分布**, 是指满足如下的可加性: 若 U 是有限个彼此不相交紧开子集 $U_i(1\leqslant i\leqslant n)$ 的并集, 则 $\mu(U)=\sum_{i=1}^{n}\mu(U_i)$.

X 上的 p-adic 分布叫作是 p-adic **测度**, 是指存在实数 B, 使得对 X 的每个紧开子集 U, 均有 $|\mu(U)|_p\leqslant B$.

为了验证一个 p-adic 分布 μ 是否为测度, 只需对 X 中的每个区间 $I=a+(p^N)$ 验证是否 $|\mu(I)|_p\leqslant B$ 即可, 因为由引理 10, 紧开子集 U 是有限个彼此不相交的区间 $I_i(1\leqslant i\leqslant n)$ 的并. 于是若 $|\mu(I_i)|_p\leqslant B(1\leqslant i\leqslant n)$, 则由 μ 是 p-adic 分布以及 p-adic 赋值特性知

$$|\mu(U)|_p=\left|\sum_{i=1}^{n}\mu(I_i)\right|_p\leqslant\max_{1\leqslant i\leqslant n}\{|\mu(I_i)|_p\}\leqslant B.$$

进而, N 级区间 $a+(p^N)$ 是 p 个彼此不相交的 $N+1$ 级区间 $a+jp^N+(p^{N+1})(0\leqslant j\leqslant p-1)$ 的并集. 如果 μ 满足可加性, 则

$$\mu(a+(p^N))=\sum_{j=1}^{p-1}\mu(a+jp^N+(p^{N+1})). \qquad (*)$$

反过来，利用引理 10 不难证明，若(*)式对每个区间 $a+(p^N)$ 均成立，则 μ 满足一般的可加性. 所以为证 μ 为 p-adic 分布，只需验证(*)即可.

例 1 取 $X=\mathbf{Z}_p$，定义 $\mu(a+(p^N))=p^{-N}\in\mathbf{Q}_p$. 易知(*)式成立，这个 p-adic 分布表示成 μ_H(Haar 分布). 当 $N\to+\infty$ 时，$|\mu_H(a+(p^N))|_p=|p^{-N}|_p=p^N\to+\infty$，从而 μ_H 不是 p-adic 测度.

例 2 取 $\mathbf{Z}_p$ 中一个固定元素 a，定义

$$\mu(U)=\begin{cases}1, & \text{若 } a\in U,\\ 0, & \text{否则}.\end{cases}$$

易知这是 $\mathbf{Z}_p$ 的 p-adic 测度，叫 Dirac 测度.

例 3 （贝努利测度） 我们知道贝努利数 B_k 定义为 $\frac{x}{e^x-1}=\sum_{k=0}^{\infty}\frac{B_k}{k!}x^k$，其中 $B_k\in\mathbf{Q}\subset\mathbf{Q}_p$. 现在由下式定义贝努利多项式 $B_k(t)\in\mathbf{Q}[t]\subset\mathbf{Q}_p[t]$，

$$\frac{xe^{tx}}{e^x-1}=\sum_{k=0}^{\infty}\frac{B_k(t)}{k!}x^k.$$

例如 $B_0(t)=1,B_1(t)=t-\frac{1}{2},B_2(t)=t^2-t+\frac{1}{6},B_3(t)=t^3-\frac{3}{2}t^2+\frac{1}{2}t,\cdots$. 易知 $B_k(0)=B_k$. 现在对每个区间 $a+(p^N)(a\in\mathbf{Z},0\leqslant a<p^N,N\geqslant 0)$定义

$$\mu_{B,k}(a+(p^N))=p^{N(k-1)}B_k\left(\frac{a}{p^N}\right)\in\mathbf{Q}\subset\mathbf{Q}_p$$
$$(k=0,1,2,\cdots).$$

这里 $B_k\left(\frac{a}{p^N}\right)$ 是 $B_k(t)$在 $t=\frac{a}{p^N}$处的值. 经过初等运算可知对每个 $k\geqslant 0$，$\mu_{B,k}$满足(*)式，所以是 p-adic 分布，叫作贝努利分布. 例如对 $N\geqslant 0,0\leqslant a<p^N$，

$$\mu_{B,0}(a+(p^N))=p^{-N}(\text{Haar 分布}),$$

$$\mu_{B,1}(a+(p^N))=B_1\left(\frac{a}{p^N}\right)=\frac{a}{p^N}-\frac{1}{2},$$

$$\mu_{B,2}(a+(p^N))=p^N B_2\left(\frac{a}{p^N}\right)=p^N\left(\frac{a^2}{p^{2N}}-\frac{a}{p^N}+\frac{1}{6}\right).$$

所有的 $\mu_{B,k}(k\geqslant 0)$都不是 p-adic 测度，但是用它们可构作出一批 p-adic 测度来. 首先注意：若 μ_1 和 μ_2 是 X 上的 p-adic 分布，则对 $\alpha_1,\alpha_2\in\boldsymbol{Q}_p$，$\alpha_1\mu_1+\alpha_2\mu_2$ 也是 p-adic 分布，其中$(\alpha_1\mu_1+\alpha_2\mu_2)(U)=\alpha_1\mu_1(U)+\alpha_2\mu_2(U)$.

现在设 $k\geqslant 1$，$\alpha\in\mathbf{Z}$，使得 $\alpha\neq 1$ 并且 $\alpha\not\equiv 0(\bmod p)$. 对于 $\mathbf{Z}_p$ 的每个紧开子集 U，则 αU 也是紧开子集. 定义 p-adic 分布

$$\mu_{k,\alpha}(U)=\mu_{B,k}(U)-\alpha^{-k}\mu_{B,k}(\alpha U).$$

经过一系列 p-adic 计算，可证 $\mu_{k,\alpha}(k\geqslant 1, 1\neq\alpha\in\mathbf{Z}, p\nmid\alpha)$均是 p-adic 测度，叫作**贝努利测度**.

有了 X 上的 p-adic 测度 μ，便可定义 X 上连续函数对 μ 的 p-adic 积分. 对每个连续函数 $f:X\to\boldsymbol{Q}_p$ 作"黎曼和"

$$\sum_{\substack{0\leqslant a<p^N\\ a+(p^N)\subseteq X}} f(x_{a,N})\mu(a+(p^N)),$$

其中 $x_{a,N}\in a+(p^N)$. 就像通常实积分情形一样，利用 μ 是 p-adic 测度，即 $|\mu(a+(p^N))|_p\leqslant B$，可证当 $N\to+\infty$时，上面求和有 p-adic 极限，并且此极限不依赖点 $x_{a,N}$的选取方式. 称此极限为 f 在 X 上对 μ 的 p-adic 积分，表示成$\int_X f\mu$.

对所有贝努利测度 $\mu_{k,\alpha}(k\geqslant 1, 1\neq\alpha\in\mathbf{Z}\ p\nmid\alpha)$都可作 p-adic 积分，它们之间有如下的关系：对 $\mathbf{Z}_p$ 的每个紧开子集 X，

$$\int_X\mu_{k,\alpha}=k\int_X x^{k-1}\mu_{1,\alpha}.$$

这表示 $\mu_{k,\alpha}$和 $\mu_{1,\alpha}$的关系与实测度 $\mathrm{d}(x^k)$和 $\mathrm{d}x$ 之间的关系相仿. 于是对 $k\geqslant 1$，

$$\mu_{B,k}(\mathbf{Z}_p)=\mu_{B,k}(0+(p^0))=B_k(0)=B_k,$$

$$\mu_{B,k}(p\mathbf{Z}_p)=\mu_{B,k}(0+(p^1))=p^{k-1}B_k,$$

所以对 $U_p=\mathbf{Z}_p-p\mathbf{Z}_p$，则有

$$\mu_{B,k}(U_p)=\mu_{B,k}(\mathbf{Z}_p)-\mu_{B,k}(p\mathbf{Z}_p)=(1-p^{k-1})B_k.$$

现在取 $1\neq\alpha\in\mathbf{Z},p\nmid\alpha$. 则 $\alpha U_p=U_p$. 于是

$$\mu_{B,\alpha}(U_p)=\mu_{B,k}(U_p)-\alpha^{-k}\mu_{B,k}(\alpha U_p)$$
$$=(1-\alpha^{-k})(1-p^{k-1})B_k.$$

由于

$$\mu_{k,\alpha}(U_p)=\int_{U_p}\mu_{k,\alpha}=k\int_{U_p}x^{k-1}\mu_{1,\alpha},$$

所以

$$\frac{1}{\alpha^{-k}-1}\int_{U_p}x^{k-1}\mu_{1,\alpha}=(1-p^{k-1})\left(-\frac{B_k}{k}\right).$$

我们看到，上式右边恰好是本节前面定义的 p-adic zeta 函数 $\zeta_p^{(a)}(s)$在负整数处的取值. 由此可给出 $\zeta_p^{(a)}(s)$的 p-adic 积分表达式：对 $s\in\mathbf{Z}_p$，

$$\zeta_p^{(a)}(1-2a-(p-1)s)$$
$$=(\alpha^{-(2a+(p-1)s)}-1)^{-1}\int_{U_p}x^{2a+(p-1)s-1}\mu_{1,\alpha}.$$

用类似方法，可以得到数域和椭圆函数等许多 p-adic L-函数的 p-adic 积分表达式，这需要找到适当的 p-adic 测度. 所以 p-adic 积分已成为研究数论的重要工具.

参 考 文 献

[1] H. Hasse, Uber die Klassenzahl abelscher Zahlkörper, 1952. 再版于 1985.

[2] K. Iwasawa, Lectures on p-adic L-functions, Princeton Univ. Press, 1972.

[3] N. Koblitz p-adic Numbers, p-adic Analysis and p-adic zeta Functions, GTM 58, Springer-verlag, 1977.

[4] S. Lang, Cyclotomic Fields I and Ⅱ, GTM 121, 1990.

[5] L. Washington, *Introduction to* Cyclotomic Fields (第二版), GTM83, 1997.

§3.8 近代分圆域理论(Ⅱ)：有限群表示论的应用

研究群结构的一个重要手段是通过群同态将不同群加以比较. 人们希望用某类群作为“样板”而研究所有群到它们的同态. 这类群要足够一般，即其结构有充分代表性；同时又很具体，即对它们有充分的了解. 人们选择了两类群作为样板，一种为置换群. 另一种为矩阵群. 群 G 到置换群的同态叫 G 的**置换表示**，而 G 到某个 n 阶线性群 $GL_n(R)$ 的同态(其中 R 为交换环，$GL_n(R)$ 表示 R 上 n 阶可逆方阵群)叫作 G 在 R 上的 n 次**线性表示**. 本节所说群表示均指线性表示. 当 R 为复数域 $\boldsymbol{C}$ 时，叫作 G 的**复表示**. 当 R 为 C_p 或它的子域(如 $\boldsymbol{Q}_p$)和子域的整数环(如 $\mathbf{Z}_p$)时，叫作 G 的 **p-adic 表示**.

有限群的(复)表示理论起源于19世纪末期 Frobenius 的工作. 在20世纪30年代交换代数产生之后，群表示论逐渐采用模论的语言. 设 $\rho: G\rightarrow GL_n(R)$ 是有限群 G 在交换环 R 上的一个 n 次表示，则通过 ρ 将 G 作用到表示空间 R^n 上，即对 $g\in G$，g 在 R^n 上的

作用是指线性变换 $\rho(g)$. 于是群环 $R[G]$作用在 R^n 上,即 R^n 是 $R[G]$-模. 更一般地,对于任意 $R[G]$-模 M,都可采用表示论的方法研究 M 的 $R[G]$-模结构. 例如设 G 是有限阿贝尔乘法群,$\hat{G}$ 是 G 的特征群(每个特征 $\chi: G \to \boldsymbol{C}^\times$ 都是 G 的一次复表示,$\boldsymbol{C}^\times = GL_1(\boldsymbol{C})$),则 $\boldsymbol{C}[G]$中 $|G|$个元素

$$\varepsilon_\chi = \frac{1}{|G|}\sum_{g\in G}\chi(a)a^{-1} \qquad (\chi \in \hat{G})$$

是 $\boldsymbol{C}[G]$的正交幂等元,即 $\varepsilon_\chi^2=\varepsilon_\chi, \varepsilon_\chi\varepsilon_\lambda=0(\chi,\lambda\in\hat{G},\chi \neq \lambda)$. 并且 $1 = \sum_{\chi\in\hat{G}}\varepsilon_\chi$. 于是环 $\boldsymbol{C}[G]$有子模直和分解

$$R=\boldsymbol{C}[G]=\bigoplus_{\chi\in\hat{G}}R_\chi,\ R_\chi=\varepsilon_\chi R.$$

从而每个 R-模 M 均有 R-子模直和分解

$$M=\bigoplus_{\chi\in\hat{G}}M_\chi, M_\chi=\varepsilon_\chi M.$$

又如对 $p-1$ 阶循环群 G(p 为素数),则对 G 的特征 χ,$\chi(g)(g\in G)$均为 $p-1$ 次单位根. 由于 $\boldsymbol{Z}_p$ 中存在 $p-1$ 次本原单位根,所以 χ 可看成 p-adic 特征,即 $\chi: G\to\boldsymbol{Z}_p$,从而 $\varepsilon_\chi = \frac{1}{p-1}\sum_{g\in G}\chi(g)g^{-1} \in \boldsymbol{Z}_p[G]$ (注意$\frac{1}{p-1}\in\boldsymbol{Z}_p$). 于是对每个 $\boldsymbol{Z}_p[G]$-模 M 均有 $\boldsymbol{Z}_p[G]$-子模直和分解

$$M=\bigoplus_{\chi\in\hat{G}}M_\chi, M_\chi=\varepsilon_\chi M.$$

数论中采用表示论可以上溯到狄里赫利甚至高斯. 二次剩余的勒让得符号,高斯和以及狄里赫利 L-函数 $L(s,\chi)$中使用的模 m 狄里赫利特征就是有限阿贝尔群$(\boldsymbol{Z}/m\boldsymbol{Z})^\times$ 的特征(一次表示). 他们在研究高斯和以及 $L(s,\chi)$的性质中均使用了特征的正交关系. 对于数域的伽罗华扩张 L/K,伽罗华群 $G=\mathrm{Gal}(L/K)$自然地作用在 L 的许多数论对象上(如整数环 O_L,L 的分式理想群 $I(L)$,理想类群 $C(L)$和单位群 $U(L)$等). 这些对象作为交换群也均为 $\boldsymbol{Z}$-模,从而均为 $\boldsymbol{Z}[G]$-模. 研究这些对象的 $\boldsymbol{Z}[G]$-模结构(伽罗华

模结构)是代数数论的重要课题.例如对伽罗华扩张 $K/\boldsymbol{Q}$,$G=\mathrm{Gal}(K/\boldsymbol{Q})$,则 O_K 是循环 $\boldsymbol{Z}[G]$-模相当于说存在 $\alpha\in O_K$,使得$\{\sigma(\alpha)\mid\sigma\in G\}$是 O_K 的一组整基.这类研究一直进行到现在并引发出许多深刻的结果和猜想.在分圆域理论中主要涉及类群或单位群的 $\boldsymbol{Z}[G]$-模或 $\boldsymbol{Z}_p[G]$-模结构问题,其中 G 是有限阿贝尔群.第一个重要结果是 Stickelberger 于 1890 年得到的.

设 K 是阿贝尔域,则它是分圆域 $\boldsymbol{Q}(\zeta_m)$的子域,其中 m 为域 K 的导子.伽罗华群 $G=\mathrm{Gal}(K/\boldsymbol{Q})$同构于$(\boldsymbol{Z}/m\boldsymbol{Z})^x$ 的一个商群.对$(a,m)=1$,$\boldsymbol{Q}(\zeta_m)$中自同构 $\sigma_a(\sigma_a(\zeta_m)=\zeta_m^a)$在 K 中的限制仍记为 σ_a.元素

$$\theta=\theta(K)=\sum_{\substack{1\leqslant a<m\\(a,m)=1}}\frac{a}{m}\sigma_a^{-1}\in\boldsymbol{Q}[G]$$

现在叫作 Stickelberger 元素,而 $I=\boldsymbol{Z}[G]\wedge\theta\boldsymbol{Z}[G]$是 $\boldsymbol{Z}[G]$的理想,叫作 Stickelberger 理想.

Stickelberger **定理**(1890 年) 对每个 $\beta\in I$ 和 K 的每个分式理想 A,β 作用 A 得到的 A^β 必为主分式理想.换句话说,Stickelberger 理想 I 零化 K 的理想类群.

证明大意 首先分析了理想 I 的结构:对每个与 m 互素的整数 C,易证$(C-\sigma_C)\theta\in\boldsymbol{Z}[G]$.由定义即知$(C-\sigma_C)\theta\in I$.进而证明了 I 是所有这种元素$(C-\sigma_C)\theta$ 生成的理想.然后考虑高斯和的素理想分解.对于属于分圆域 K 的各种高斯和 g,Stikelberger 明显地给出主理想(g)在 O_K 中的素理想分解式

$$(g)=gO_K=P_1^{a_1}P_2^{a_2}\cdots P_r^{ar},$$

其中 $G=\mathrm{GaL}(K/\boldsymbol{Q})$在素理想集合$\{P_1,\cdots,P_r\}$上的作用是可迁的,即存在素理想 P,使 $P_i=P^{\sigma_i}$(对某个 $\sigma_i\in G$).于是$(g)=P^\beta$,其中 $\beta=a_1\sigma_1+\cdots+a_r\sigma_r\in\boldsymbol{Z}[G]$.具体的计算表明 β 有形式$(C-\sigma_C)\theta$.β 把 P 变成主理想(g),即是说 β 把理想类$[P]$零化.使用

足够多的高斯和的素理想分解式，便得到所有元素$(C-\sigma_C)\theta$把每个理想类零化，于是I零化K的整个理想类群.

1932年，法国人Herbrand用上述结果推进了库默尔关于分圆域的工作. 对于素数$p\geqslant 5$，以h_p和h_p^+分别表示$K=\boldsymbol{Q}(\zeta_p)$和$K^+=\boldsymbol{Q}(\zeta_p+\zeta_p^-)$的理想类数. 库默尔证明了(见§2.2所述)：$h_p^-=h_p/h_p^+$是正整数，并且$p|h_p\Leftrightarrow p|h_p^-\Leftrightarrow p$除尽贝努利数$B_2$，$B_4$，$\cdots,B_{p-3}$某个的分子.

$$\left(\text{即存在 } n,\ 1\leqslant n\leqslant\frac{p-3}{2},\ \text{使得 } v_p(B_{2n})\geqslant 1.\right)$$

现在采取表示论观点. 以A和A^+分别表示K和K^+的理想类群$C(K)$和$C(K^+)$的Sylow p-子群. 则$|A|$和$|A^+|$分别等于h_p和h_p^+的p-部分. 于是

$$p|h_p\Leftrightarrow|A|>1\Leftrightarrow A\neq(1).$$

群$G=Gal(K/\boldsymbol{Q})$自然作用于类群$C(K)$之上，也自然作用于Sylow子群A之上. 另一方面，A是有限阿贝尔p-群. 每个元素$a\in A$的阶有形式$p^l(l\geqslant 0)$. 对每个$\boldsymbol{Z}_p$中p-adic整数$\alpha=\sum\limits_{i=0}^{\infty}a_ip^i$，当$i\geqslant l$时$a^{p^i}=1$. 因此可认为$\alpha$对$a$的作用为$a^\alpha=a^{a_0+a_1p+\cdots+a_{l-1}p^{l-1}}$. 于是$\boldsymbol{Z}_p$自然作用于$A$上，从而$A$是$\boldsymbol{Z}_p[G]$-模. G是$p-1$阶循环群(同构于$(\boldsymbol{Z}/p\boldsymbol{Z})^\times$)，所以$G$的特征可看成是模$p$的狄里赫利特征，并且取值于$\boldsymbol{Z}_p$. $\hat{G}$为$p-1$阶循环群是由Techmüller特征w生成的，其中对每个与p互素的整数a，$w(a)$为$\boldsymbol{Z}_p$中的$p-1$次单位根，使得$w(a)\equiv a(\mathrm{mod}\ p)$. 于是$\hat{G}=\{w^i|0\leqslant i\leqslant p-2\}$，以$\varepsilon_i$简记前面定义的幂等元$\varepsilon_{w^i}$，则$A$作为$\boldsymbol{Z}_p[G]$-模有直和分解

$$A=\bigoplus_{i=0}^{p-2}A_i,\quad A_i=\varepsilon_iA.$$

注意偶特征全体$\{w^i|0\leqslant i\leqslant p-2,\ 2|i\}$是$\mathrm{Gal}(K^+/\boldsymbol{Q})$的特征群，

由此可证$\bigoplus_{\substack{i=0\\2|i}}^{p-2}A_i$同构于$A^+$．于是对

$$A^- = \bigoplus_{\substack{i=0\\2\nmid i}}^{p-2}A_i = A_1 \oplus A_3 \oplus A_5 \oplus \cdots \oplus A_{p-2}.$$

可知$|A^-|$即为h_p^-的p-部分，即$|A^-|=p^s$，其中$s=v_p(h_p^-)$．容易证明$A_1=(1)$．因此库默尔结果可以说成：

$p|h_p^-$（即对某个$i=3, 5, \cdots, p-2$，使$A_i \neq (1)$）

$\Leftrightarrow$存在某个$i=2, 4, \cdots, p-3$，使$v_p(B_i) \geqslant 1$．

1932年，Herbrand证明了：对每个$2\nmid i$，$3 \leqslant i \leqslant p-2$，若$A_i \neq (1)$，则$v_p(B_{p-i}) \geqslant 1$．1976年，Ribet证明了它的逆：若$v_p(B_{p-i}) \geqslant 1$，则$A_i \neq (1)$．于是：$A_i \neq (1) \Leftrightarrow v_p(B_{p-i}) \geqslant 1$．这就细化了库默尔的上述结果．Herbrand结果的证明使用Stickelberger定理，在发明了p-adic$L$-函数之后，可以对证明有更好的理解，而Ribet结果的证明则使用了深刻的模曲线和代数几何手段．1987年，K. Rubin采用Kolyvagin更为初等的方法证明了Ribet结果．1992年，Harder和Pink给出Ribet结果的另一个证明．

1958年，Hasse的学生Leopoldt把库默尔的结果$p|h_p^+ \Rightarrow p|h_p^-$（即$p|h_p \Leftrightarrow p|h_p^-$）作了同样风格的细化．

Leopoldt **反射定理**　设i为偶数，$0 \leqslant i \leqslant p-3$，而$j=p-2-i$．则$A_i$的$p$-秩$\leqslant A_j$的$p$-秩$\leqslant A_i$的$p$-秩$+1$．这里对每个有限阿贝尔$p$-群$A$，$A$的$p$-秩是指$A/A^p$在域$\boldsymbol{F}_p$上的维数．

这个定理可以推广到阿贝尔数域上，用p-adic L-函数可以给出反射定理的一种解释：作为Mazur和Wiles于1984年所证的岩泽主猜想（对$\boldsymbol{Q}$）的推论，可以得到群同构：

若$i \in \{3, 5, \cdots, p-2\}$，则$A_i \cong \boldsymbol{Z}_p / L_p(0, \omega^{1-i})\boldsymbol{Z}_p$.

而反射定理可由p-adic$L$-函数的解析特性推出．

库默尔在给Kronecker的信中把$p \nmid h_p^+$（对每个奇素数p）

称作“待证明的定理”. 现在通常称作 Vandiver 猜想. 目前对于 $p<4\times 10^6$ 均已验证这个猜想是正确的，并且由 Vandiver 猜想可以推出分圆域理论中不少重要的结论，但这个猜想至今未解决.

以上讲述了 20 世纪对于分圆域(和阿贝尔域)的类群所作的研究. 除了上述之外，综合使用代数和解析方法，1971 年 Montgomery 和 Uchida 各自独立地证明了库默尔的如下猜想：对于奇素数 p，$h_p=1$ 当且仅当 $p\leqslant 19$. 不久，Masley 于 1976 年又采用同样方法决定出类数为 1 的全部分圆域 $\boldsymbol{Q}(\zeta_m)$(共 30 个). 下面讲述对于分圆域单位群的研究结果. 对每个正整数 $m\not\equiv 2(\mathrm{mod}4)$，分圆域 $\boldsymbol{Q}(\zeta_m)$的(整数环 $O_K=\boldsymbol{Z}[\zeta_m]$的)单位群 U_m 有直和分解

$$U_m=W_m\times V_m,$$

其中 W_m 是 O_K 中单位根循环群，阶为 m(若 $4|m$)或 $2m$(若 $2\nmid m$)，而 V_m 是秩 $r=\frac{1}{2}\varphi(m)-1$ 的自由阿贝尔群. 对于 K 的最大实子域 $K^+=\boldsymbol{Q}(\zeta_m+\bar{\zeta}_m)$，其单位群为 $U_m^+=(\pm 1)\times V_m^+$，其中 V_m^+ 也是秩 $r=\frac{1}{2}\varphi(m)-1$ 的自由阿贝尔群. 库默尔证明了 U_m 中单位均可表成 U_m^+ 中单位(实单位)和单位根之积，所以可以取 V_m 为 V_m^+. 即可取 K^+的 r 个实单位，同时是 K 和 K^+的基本单位系. 当 m 很大时，要具体构作这样一组实单位系是很困难的. 但是库默尔发现 r 个非常简单的实单位($\zeta=\zeta_m$)，

$$\varepsilon_a=\frac{1-\zeta^a}{1-\zeta}\zeta^{\frac{1-a}{2}}=\frac{\zeta^{-a/2}-\zeta^{a/2}}{\zeta^{-1/2}-\zeta^{1/2}}=\frac{\sin\frac{a\pi}{m}}{\sin\frac{\pi}{m}}\left(\begin{array}{l}1<a<m/2,\\(a,\ m)=1.\end{array}\right)$$

称作**分圆单位**. 对于 $m=p$ 的情形，他证明这 r 个分圆单位是独立的(即它们之间没有乘法关系)，从而生成 U_p 和 U_p^+ 的一个秩 r 自由阿贝尔子群 C_p^+，叫作**分圆单位群**. 于是 U_p^+/C_p^+ 和 U_p/C_p^+ 都是有限群. 由于 K^+的理想类数 h_p^+ 的公式中包含这些分圆单位的

对数值，库默尔由此证明出当 $m=p$ 时

$$[U_p : W_pC_p^+]=[U_p^+ : C_p^+]=h_p^+.$$

对于 $m=p^n$ 有类似的结论，但是希尔伯特在《数论报告》中发表看法，认为对一般的 m，库默尔构作的 r 个分圆单位可能不独立，从而分圆单位群 C_m^+ 的秩会小于整个单位群 U_m(和 U_m^+)的秩 $r=\frac{1}{2}\varphi(m)-1$. 1965 年，Ramachandra 在数学上证明了希尔伯特论断的正确性. 比如说，当 m 有 4 个不同素因子时，C_m^+ 的秩一定小于 r. 1978 年，Sinnott 把分圆域 $\boldsymbol{Q}(S_m)$和它所有阿贝尔子域的分圆单位放在一起生成一个更大的分圆单位群$\widetilde{C}_m^+$，证明这个群的秩为 r，并且对它的伽罗华模结构作了细致的上同调计算，得到 $[U_m^+ : \widetilde{C}_m^+]=2^b\cdot h_m^+$，其中 b 为某个确定的非负整数. 1980 年，Sinnott 又把此结果推广到任意阿贝尔数域.

上述结果体现出实阿贝尔域的单位群和理想类数之间的联系，人们自然要问:经典代数数论的两个最主要的研究对象——类群和单位群，至少对分圆域的情形是否在结构上有联系？比如对 $K^+=\boldsymbol{Q}(\zeta_p+\bar{\zeta}_p)$，库默尔的公式$[U_p^+ : C_p^+]=h_p^+$ 表明群 U_p^+/C_p^+ 和 K^+ 的理想类群 $C(K^+)$有相同的阶. 进而要问，它们作为伽罗华模(即作为 $\boldsymbol{Z}[G]$-模，$G=\mathrm{Gal}(K^+/\boldsymbol{Q})\cong(\boldsymbol{Z}/p\boldsymbol{Z})^x/(\pm1)$)是否同构?有反例($p=62501$)表明这并不永远成立. 于是退一步问:它们作为有限阿贝尔群是否同构？目前这个问题没有结论. 甚至这两个群的 p-部分作为阿贝尔群是否同构也不知道. 根据上述，$C(K^-)$的 Sylow$p$-子群为 $A^-=\bigoplus_{\substack{i=2\\2|i}}^{p-3}A_i\,(A_i=\varepsilon_iA^+)$，而 U_p^+/C_p^+ 的 Sylow$p$-子群(表示成 $U^{(p)}$)作为 $\boldsymbol{Z}_p[G]$-模也有分解

$$U^{(p)}=\bigoplus_{\substack{i=2\\2|i}}^{p-3}U_i^{(p)},\ U_i^{(p)}=\varepsilon_iU^{(p)}.$$

目前得到的最好结果是对每个 $2\leqslant i\leqslant p-3$，$2|i$，均有 $|A_i|=$

$|U_i^{(p)}|$，这是 Mazur-Wiles 证明的岩泽主猜想(对 $\boldsymbol{Q}$)的一个推论. 1990 年 Kolyvagin 用比较初等的方法(Euler system)重新证明了这件事，并且 Rubin 用 Kolyvagin 方法重新证明了对 $\boldsymbol{Q}$ 的岩泽主猜想.

h_p^+ 是类群 $C(K^+)$的阶，库默尔说它也是群 U_p^+/C_p^+ 的阶. 但是相对类数 h_p^- 却只是两个类数的商 h_p/h_p^+，它是否也有别的代数刻划方式?1962 年，岩泽对于 $m=p^l$ 的情形给出 $h_{p^l}^-$的一种代数刻划. 令 $K=\boldsymbol{Q}(\zeta_{p^l})$，$G=Gal(K/\boldsymbol{Q})=\{\sigma_a \mid 1\leqslant a\leqslant p^l, (a, p)=1\}$. 我们有 Stickelberger 元素

$$\theta=\frac{1}{p^l}\sum_{\substack{a=1\\(a,p)=1}}^{p^l} a\sigma_a^{-1}\in\boldsymbol{Q}[G]$$

和 $R=\mathbf{Z}[G]$的 Stickelberger 理想 $I=R\cap\theta R$. 考虑 R 和 I 的“负”部分.

$$R^-=(1-\sigma_{-1})R=\{x\in R \mid \sigma_{-1}x=-x\},\ I^-=I\cap R^-,$$

则 I^-是 R^-的 R-子模. 岩泽证明了:$[R^-:I^-]=h_{p^l}^-$.
1978 和 1980 年，Sinnott 把这个结果推广到任意分圆域和阿贝尔域上. 以 C 和 C^+分别表示 $K=\boldsymbol{Q}(\zeta_{p^l})$和 K^+的理想类群，则 $h_{p^l}^-=[C:C^+]$. 自然要问是否有 $\mathbf{Z}[G]$-模同构 $R^-/I^-\cong C/C^+$? 这同样有反例($p=4027$). 人们又退而考虑它们的 p-部分. C/C^+的 Sylow p-子群就是前面的 A^-. 记$(R^-/I^-)_p$ 为 R^-/I^-的 p-部分，那么作为 $\mathbf{Z}_p[G]$-模是否 $A^-\cong(R^-/I^-)_p$?目前只证明在 Vandiver 猜想 $p\nmid h_p^+$ 成立时这是对的.

以上是研究阿贝尔域的类群和单位群的伽罗华模性质. 1920 年，Hecke 研究任意数域的阿贝尔扩张 L/K，即 Gal(L/K)为阿贝尔群. 他用类域论构作了一种 L-函数 $L(s, \chi)$，给出解析延拓和函数方程(叫作 Hecke L-函数). 利用它的解析特性得到 O_K 中(在 L 中具有某种分解特性)素理想的分布规律. 1923 年，Artin

对任意伽罗华扩张 L/K 构作了一种 L-函数，使用伽罗华群 $G=\mathrm{Gal}(L/K)$的复表示. Artin L-函数一开始用来研究 O_K 中(在 L 中具有某种分解特性)素理想的分布，后来用于研究 L 的单位群.

我们介绍 Artin L-函数的定义. 设 $\rho: G\to GL_n(\boldsymbol{C})$是 G 的一个复表示，χ 为表示 ρ 的特征. 对于 O_K 和 O_L 的非零素理想 $\mathscr{P}$ 和 P，$P|\mathscr{P}$，以 Z_P 和 T_P 分别表示 P 对于 L/K 的分解群和惰性群. 令 V_P 为表示空间 $V=\boldsymbol{C}^n$ 对 T_P 的固定子空间，即

$$V_P=\{v\in V\,|\,gv=v(对每个\ g\in T_P\},$$

则 V_P 是 $\boldsymbol{C}[Z_P/T_P]$-模. 当 P 不分歧时 $T_P=\{1\}$，从而对几乎每个 P 均有 $V_P=V$. 令 σ_P 为 Z_P/T_P 的 Frobenius 自同构生成元，则 $1-N\mathscr{P}^{-s}\sigma_P$ 作用于 V_P 上，并且它在 V_p 上这种作用的行列式 $\det(1-N\mathscr{P}^{-s}\sigma_P)$只依赖于 $\mathscr{P}$，是 $N\mathscr{P}^{-s}$的多项式(次数$\leqslant n=[l:K]$)，它也只和表示 ρ 的共轭类有关，从而是特征 χ 的函数. Artin$L$-函数定义为

$$L(s,\chi)=\prod_{\mathscr{P}}L_{\mathscr{P}}(s,\chi),\ L_{\mathscr{P}}(s,\chi)=\det(1-N_{\mathscr{P}^{-s}}\sigma_p)^{-1}.$$

可以证明它在区域 $\mathrm{Re}s>1$ 中收敛，并且有如下的基本性质：

(1) 若 χ_1 和 χ_2 均为 G 的特征，则

$$L(s,\chi_1+\chi_2)=L(s,\chi_1)L(s,\chi_2);$$

(2) 若 H 是 G 的正规子群，χ 为 G/H 的特征. 将 χ 提升成 G 的特征 χ'，则 $L(s,\chi')=L(s,\chi)$；

(3) 若 H 为 G 的子群，χ 为 H 的特征. 以 $\mathrm{ind}_H^G\chi$ 表示 χ 在 G 上的诱导特征，则 $L(s,\mathrm{ind}_H^G\chi)=L(s,\chi)$；

(4) 若 G 是阿贝尔群，χ 是 G 的一次表示特征，则 $L(s,\chi)$ 本质上是 Hecke$L$-函数，所以有解析延拓与函数方程，并且它只在 $s=1$ 有极点，且是单极点.

根据 Brauer 定理，G 的每个特征都是循环群 H_i 的特征 χ_i 所诱导的特征的线性组合：$\chi=\sum_{i=1}^{m}n_i\mathrm{ind}_{H_i}^G\chi_i(n_i\in\boldsymbol{Z})$，所以 $L(s,\chi)$

$= \prod_{i=1}^{m} L(s, \chi_i)^{n_i}$，而由(4)知 Artin L-函数均可解析延拓成整个复平面上的亚纯函数. Artin 的一个著名猜想为：

若 χ 不包含平凡特征(取值恒为 1)，则 $L(s, \chi)$为复平面上全纯函数.

由(3)和(4)可知，由 1 次表示特征诱导的 χ，$L(s, \chi)$是全纯的. 由 Artin 猜想可推出代数数论中关于狄德全 zeta 函数的重要猜想:对任意数域扩张 M/N，$\zeta_M(s)/\zeta_N(s)$是复平面上全纯函数. 这个猜想对许多情形得到证明，但未完全解决.

以上我们概括地介绍了有限群表示理论在代数数论中的某些应用. 下一个发展是无限群表示理论在数论中的应用，它开创了代数数论的一个新篇章:现代代数数论.

参 考 文 献

(包括上节参考文献中的[4]和[5])

[1] V. Kolyvagin, Euler Systems, The Grothendieck Festschrift, Vol. Ⅱ, p. 435～483. Birkhauser, 1990.

[2] B. Mazur and A. Wiles, class fields of abelian extensions of $\boldsymbol{Q}$, Invent. Math. 76(1984), 179～330.

[3] K. Ribet, A modular construction of unramified p-extensions of $\boldsymbol{Q}(\mu_p)$, Invent. Math. 34(1976), 151～162.

[4] K. Rubin, The "main conjectures" of Iwasawa theory for imaginary quadratic fields, Invent. Math. 103(1991), 25～

68.
[5] W. Sinnott, On the Stickelberger ideal and the circular units of an abelian fields, Invent. Math. 62(1980), 181～234.
[6] L. Stickelberger, Uber eine Verallgemeinering der Kreistheilung, Math. Ann. 37(1890), 321～367.
[7] A. Wiles, The Iwasawa conjecture for totally real fields, Ann. of Math. 3, (1990), 493～540.

第四章　现代代数数论(1967～　)

在整个 20 世纪里，代数数论一直是十分活跃和延续不断的研究领域. 严格说来，近代和现代的代数数论没有明显的分界. 我们把 1967 年看作是现代代数数论的开始，是因为这一年在代数数论研究中有两件具有重要意义的事情. 一件是韦依出版了《基础数论》(Basic Number Theory)一书. 这本书总结了 20 世纪前 70 年代数数论的重要研究方法，把这些方法加以系统化、标准化和抽象化. 韦依在书中把局部-整体原则凝聚在 Adèle 和 Idèle 语言之中，完善而统一地采用局部紧群上调和分析来讲述代数数论，给人以面目一新的感觉. 此书已成为后人学习和研究代数数论的必读经典著作，对代数数论的发展影响很大. 另一件是朗兰兹提出了一个重要的猜想，后来发展成规模宏伟的"朗兰兹纲领"，它把群的无穷维表示理论引入到代数数论中，使代数数论以及许多相关领域的研究进入一个崭新的时期. 1967 年以前，代数数论已经和代数几何、复分析、有限群表示理论和解析数论等融合交织在一起. 1967 年以后，代数数论在更高的层次上又与李群、无限群表示理论、群上调和分析和微分几何密不可分. 20 世纪后 30 年代数数论的发展还有两个显著的特点.

(1) 这是代数数论大丰收的时期. 20 世纪前 70 年代数数论取得许多重要成果. 但给人印象最深的是涌现出丰富的思想和一大批新学科. 它们所蕴育的能量在 20 世纪 70 年代以后达到总的爆发，产生出一批重大成果. 这些成果的集中体现是三项菲尔兹奖和一项沃尔夫奖. 这就是德林证明高维韦依猜想、法廷斯证明莫代尔猜想、德林费尔德证明函数域上二维朗兰兹局部猜想(这三项成果分别获得 1978，1986 和 1990 年菲尔兹奖)以及怀尔斯证

明费马猜想(怀尔斯和朗兰兹同时获得 1996 年沃尔夫奖).

(2) 20 世纪 70 年代以来，代数数论在取得上述辉煌理论成就的同时，在计算机科学和数字通信信息科学的许多领域得到重大的应用，并且常常体现为代数数论的最新成果的深刻应用. 代数数论同时扮演着皇后和仆人的角色，很好地发挥着“可上九天揽月，可下五洋捉鳖”的功能.

我们在本章中先介绍 1967 年两项重大事件的意义，然后以四项重大成就为线索介绍近 30 年代数数论的发展. 在第五章介绍代数数论的几项重要应用.

§4.1 韦依的《基础数论》(1967 年)

大数学家常常给自己所写的书起一个小小的名字，《基础数论》实际上讲述了很高深的数论. 本书共分两部分，第一部分是用局部紧群上的调和分析统一讲述整体域(代数数域和函数域)的基本理论，第二部分则是用单代数的语言统一讲述整体域上的类域理论.

韦依在书的前言中说，这本书是讨论代数数论“近 30 年”(指 1967 年以前的 30 年)的发展，也就是局部紧群和其上的测度与积分在代数数论研究中所起的作用. 他说，在狄里赫利和厄尔米特时代，甚至在闵柯夫斯基年代，数论问题中出现的“连续变量”就好像变魔术一样. 但是现在我们知道，实数域不过是与所有 p-adic 数域完全平等的大家庭成员之一. 韦依说，局部域的许多重要事实和有价值的结果可以不用局部紧群性质，而只用代数语言，从而在更一般的条件下也成立. 某些数学家似乎觉得他们完

全领悟了上帝的旨意. 但是他本人是在做数学而不是做神学，他只是对数学材料加以和谐的处理，满足逻辑和美学的要求，他在讲述类域论时没有明显使用任何上同调语言. 从这个意义上说，书的第一部分可以认为是"现代"的，而第二部分是"非现代"的. 但是，任何一位熟悉伽罗华上同调的人，都很容易把书中用单代数讲述的类域论翻译成上同调语言. 他的用意是"想对于行驶特定航线的航船设置较少的装备，以便可以承载更多的货物. 如果装备太多，船就要沉掉."

韦依所说的"特定航线"指的是代数数论和类域论. 书中讲述的多数结果都是经典的和已知的. 但是韦依采取了全新的观点和处理方法，这种观点和方法现已成为后人从事代数数论研究的样板，这是此书的最重要价值.

书中使用了由所有局部信息来反映整体域 K 信息的理想语言和工具，这就是 K 的 Adele 环和 Idele 群. 这两个概念在韦依之前已经被 Chevalley 和 Iwasawa 所采用，但是韦依在书中运用得最为充分，并且把数域和函数域统一讲述. 现在我们以数域为例介绍现代代数数论的这两个概念(函数域情形除了韦依的书之外，初学者可阅读文献中 W. -C. Winnie Li 的书).

设 K 为代数数域，以 Σ_K 表示 K 的所有素除子所成的集合. 对 K 的每个素除子 $\mathscr{P}\in\Sigma_K$，$K_{\mathscr{P}}$ 是 K 对 $\mathscr{P}$ 的局部域. 当 $\mathscr{P}$ 是阿基米德素除子时，$K_{\mathscr{P}}=\boldsymbol{R}$ 或 $\boldsymbol{C}$. 而当 $\mathscr{P}$ 是非阿素除子时，以 $O_{\mathscr{P}}$ 表示 $K_{\mathscr{P}}$ 的整数环. 所有加法群 $K_{\mathscr{P}}$ 都是局部紧群. 我们自然会想到无限直积群

$$\prod_{\mathscr{P}\in\Sigma_K} K_{\mathscr{P}}=\{(x_{\mathscr{P}})\mid x_{\mathscr{P}}\in K_{\mathscr{P}}\}.$$

它收集了 K 的所有局部信息，但是它没有满意的拓朴性质. 比如说，作为无限个局部紧群的直积，它的积拓朴不是局部紧的. 合适的对象是考虑它的子群

$$A_K=\{(x_{\mathscr{P}})\in\prod_{\mathscr{P}\in\Sigma_K} K_{\mathscr{P}}\mid \text{对几乎所有 }\mathscr{P},\ x_{\mathscr{P}}\in O_{\mathscr{P}}\}.$$

这里"对几乎所有 $\mathscr{P}$"是指"除了所有阿基米德素除子和有限个非阿素除子之外"的意思. 事实上 A_K 是一个交换环，叫作 K 的 Adele 环. 以 Σ_∞ 表示 K 的所有(有限个)阿基米德素除子组成的集合，对 Σ_K 的每个有限子集 S，并且 $S\supseteq\Sigma_\infty$，记

$$A_K(S)=\prod_{\mathscr{P}\in S}K_{\mathscr{P}}\cdot\prod_{\mathscr{P}\in\sum_K-S}O_{\mathscr{P}}.$$

这是 A_K 的子环，它是有限个局部紧拓朴空间和无限个紧拓朴空间的直积，对于积拓朴 $A_K(S)$是局部紧的. 如果 S_1 和 S_2 是上述两个素除子集合，$S_1\geqslant S_2$，则 $A_K(S_1)\geqslant A_K(S_2)$，并且 $A_K(S_2)$中的拓朴即是由 $A_K(S_1)$的拓朴诱导出来的子集拓朴. 所有 $A_K(S)$的(正向)极限就是 Adele 环 A_K. 对每个满足上述条件的 S，集合 $\prod_{\mathscr{P}\in S}W_{\mathscr{P}}\prod_{\mathscr{P}\in\sum_K-S}O_{\mathscr{P}}$ 为包含零元素 0 的开集，其中 $W_{\mathscr{P}}$ 是 $K_{\mathscr{P}}$ 中包含 0 的开集. 所有这样集合构成 0 的开集基，可知 A_K 对这样的拓朴是局部紧的. A_K 中元素$(x_{\mathscr{P}})$叫作 K 的一个 adele.

对于 K 中每个元素 x，则对几乎所有 $\mathscr{P}$ 均 $x\in O_{\mathscr{P}}$. 所以(x)(即$(x_{\mathscr{P}})$，其中对每个 $\mathscr{P}$，$x_{\mathscr{P}}=x$)是 adele. 由此可把 K 嵌到 A_K 中(这叫对角嵌入)：

$$K\subseteq A_K,\ x\mapsto(x),$$

从而 K 可看成是 A_K 的加法子群和子环. 可以证明

(1) K 是 A_K 的离散子群，并且商群 A_K/K 是紧的.

Adele 环 A_K 的单位群叫作 K 的 Idele 群，表示成 J_K. 即

$$J_K=\{(x_{\mathscr{P}})\in\prod_{\mathscr{P}\in\sum_K}K_{\mathscr{P}}^{\times}\mid\text{对几乎所有 }\mathscr{P},\ x_{\mathscr{P}}\in U_{\mathscr{P}}\}.$$

这里 $U_{\mathscr{P}}=O_{\mathscr{P}}^{\times}$ 是整数环 $O_{\mathscr{P}}$ 的单位群. 但是由 A_K 诱导出来的 J_K 中拓朴不够好，我们用更细的拓朴，即取所有 $\prod_{\mathscr{P}\in S}W_{\mathscr{P}}\cdot\prod_{\mathscr{P}\in\sum_K-S}U_{\mathscr{P}}$($S$ 过所有上述集合，$W_{\mathscr{P}}$ 为 $K_{\mathscr{P}}^{\times}$ 中包含 1 的开集)为 1 在 J_K 中的开集基，这使 J_K 为局部紧群. J_K 中元素叫作 K 的一

个 Idele. K^x 也可由对角嵌入成为 J_K 的子群，商群 J_K/K^x 叫作 K 的 Idele 类群，表示成 C_K. 以 $|\ |_{\mathscr{P}}$ 表示 $K_{\mathscr{P}}$ 的标准赋值，我们有群同态

$$|\ |: I_K \to \mathbf{R}^x_{>0}(\text{正实数乘法群}),$$

$$|(x_{\mathscr{P}})| = \prod_{\mathscr{P}\in\sum_K} |x_{\mathscr{P}}|_{\mathscr{P}}(\text{叫 Idele}(x_{\mathscr{P}})\text{的范})$$

（注意对每个 Idele$(x_{\mathscr{P}})$，对几乎所有 $\mathscr{P}$，$x_{\mathscr{P}}\in U_{\mathscr{P}}$，即 $|x_{\mathscr{P}}|_{\mathscr{P}}=1$，从而 $\prod\limits_{\mathscr{P}\in\sum_K} |x_{\mathscr{P}}|_{\mathscr{P}}$ 是有限乘积）. 这个映射的核为 I_K 的子群 $I_K^1 =\{(x_{\mathscr{P}})\in I_K \mid |(x_{\mathscr{P}})|=1\}$，并且

$$I_K \cong I_K^1 \times \mathbf{R}^x_{>0}(\text{直积}).$$

由关于赋值的乘积公式可知 $K^x\subseteq I_K^1$. 我们有

(2) K^x 是 I_K^1 的离散子群，并且 I_K^1/K^x 是紧的.

利用 Adele 环和 Idele 群的语言和它们的拓朴性质可以证明代数数论的经典结果. 比如说，对每个代数数域 K，可以证明它的理想类群同构于 J_K/K^xU，其中 $U=\prod\limits_{\mathscr{P}} U_{\mathscr{P}}$（$\mathscr{P}$ 过 K 的所有非阿素除子，而 $U_{\mathscr{P}}$ 为局部域 $K_{\mathscr{P}}$ 中的 $\mathscr{P}$-adic 单位群）. 再证明 J^K/K^xU 是紧群并且是离散的，从而为有限群. 这就由 J_K 的拓朴性质推出 K 的理想类群是有限阿贝尔群. 类似地可以证明函数域的理想类群和零次除子类群均是有限的. 用 Adele 环可以证明函数域（即代数曲线）的黎曼-洛赫定理.

现在介绍一下如何用 Idele 语言描述类域论. 设 K 为数域. $M_0=\prod P_i^{e_i}$ 为 O_K 的一个理想，M_∞是 K 的不同实素除子之积，则 $M=M_0M_\infty$叫作域 K 的一个除子. 比如对 $K=\mathbf{Q}$，它只有一个实除子∞，所以 $M=1$，∞，$5^3\cdot 17^2\cdot\infty$，$5^3\cdot 17^2$ 都是 $\mathbf{Q}$ 的除子. 对每个 $\alpha\in K^x$，我们用同余式 $\alpha\equiv 1(\bmod^* M)$表示满足以下两个条件

(i) 对 M_0 的每个素理想因子 p_i（即 $e_i\geqslant 1$），$v_{P_i}(\alpha-1)\geqslant e_i$，即 $\alpha\equiv 1(\bmod p_i^{e_i})$.

(ii) 对 M_∞ 中每个实素除子 P，$\sigma(\alpha)>0$，其中 σ 是对应于 P 的实嵌入 $\sigma: K\to \boldsymbol{R}$.

我们以 I_M 表示 K 的分式理想 $\frac{A}{B}$（A，B 均为 O_K 中与 M 互素的理想）全体形成的群，以 P_M 表示主理想 αO_K（$\alpha\equiv 1(\bmod^* M)$）形成的群. 易知 P_M 是 I_M 的子群. 商群 I_M/P_M 叫作 K 的模 M **广义类群**. 当 $M=1$ 时，这就是通常的理想类群.

以 $K=\boldsymbol{Q}$ 为例，则 $\boldsymbol{Q}$ 的除子为 n 或 $n\infty$，其中 n 是正整数. 对于 $M=n$，则

$$I_n=\left\{\frac{a}{b}\mathbf{Z}\,|\,a,\ b\in\mathbf{Z},\ b\neq 0,\ (a,\ n)=(b,\ n)=1\right\},$$

$$P_n=\left\{\frac{a}{b}\mathbf{Z}\,|\,\frac{a}{b}\equiv 1,\ (\bmod\ n)\right\}.$$

由于 $\frac{a}{b}$ 和 $-\frac{a}{b}$ 均可生成 I_n 中同样的分式理想. 可知 $I_n/P_n\cong(\mathbf{Z}/n\mathbf{Z})^x/(\pm 1)$. 另一方面，对于 $M=n\infty$，则 $I_{n\infty}=I_n$，但是 $P_{n\infty}=\left\{\frac{a}{b}\mathbf{Z}\,|\,\frac{a}{b}\equiv 1(\bmod\ n),\ \frac{a}{b}>0\right\}$. 可知 $I_{n\infty}/P_{n\infty}\cong(\mathbf{Z}/n\mathbf{Z})^x$. 我们知道，对 $\boldsymbol{Q}$ 的分圆扩域 $\boldsymbol{Q}(\zeta_n)$，其伽罗华群 $\mathrm{Gal}(\boldsymbol{Q}(\zeta_n)/\boldsymbol{Q})$ 恰好同构于模 $n\infty$ 的广义理想类群 $I_{n\infty}/P_{n\infty}$（即 $(\mathbf{Z}/n\mathbf{Z})^x$），并且素除子 p 在 $\boldsymbol{Q}(\zeta_n)$ 中分歧，当且仅当 $p|n\infty$（K 的实除子 P 在阿贝尔扩域 L 中分歧，是指 P 在 L 中的扩充为复除子）. 由于 $\boldsymbol{Q}$ 的每个阿贝尔扩域 L 均是某个分圆域 $\boldsymbol{Q}(\zeta_n)$ 的子域，所以伽罗华群 $\mathrm{Gal}(L/\boldsymbol{Q})$ 同构于 $I_{n\infty}/P_{n\infty}$ 的某个商群 $I_{n\infty}/H$，其中 $H\supseteq P_{n\infty}$. 类似地，$\boldsymbol{Q}$ 的每个实阿贝尔扩域 L 为某个 $\boldsymbol{Q}(\zeta_n)^+$ 的子域，而 $\mathrm{Gal}(\boldsymbol{Q}(\zeta_n)^+/\boldsymbol{Q})$ 同构于 $I_n/P_n(\cong(\mathbf{Z}/n\mathbf{Z})^x/(\pm 1))$，所以 $\mathrm{Gal}(L/\boldsymbol{Q})\cong I_n/H$，其中 $H\geqslant P_n$. 以这个简单例子作为背景，我们来理解下面所述的类域论一般结果.

设 L/K 是数域的阿贝尔扩张. 对 O_K 的素理想 $\mathscr{P}$，若 $\mathscr{P}$ 在 L 中不分歧，则 Frobenius 自同构 $\left(\frac{L/K}{\mathscr{P}}\right)\in\mathrm{Gal}(L/K)$. 令 M 是

O_K 的全体在 L 中不分歧的素理想之乘积，则由此给出一个映射（群同态）

$$f=(L/K):I_M\to \mathrm{Gal}(L/K)$$

叫作 Artin 映射.

定理 1 设 L/K 是数域的阿贝尔扩张. 则存在 K 的惟一除子 f（叫扩张 L/K 的导子）满足以下二条件.

（1）K 中素除子（有限或无限）$\mathscr{P}$ 在 L 中分歧当且仅当 $\mathscr{P}|f$.

（2）对 K 的每个除子 M，$f|M$，则 Artin 映射给出同构 $I_M/H\cong \mathrm{Gal}(L/K)$，其中 $H=P_MN_{L/K}(I_M(L))$，而 $I_M(L)$ 是 L 中分子分母均与 M 互素的分式理想形成的群.

定理 2 对于 K 的每个除子 M 和 I_M 的每个子群 $H(H\geqslant P_M)$，则 K 存在惟一的阿贝尔扩张 L 使得

（1）若 K 中素除子 $\mathscr{P}$ 在 L 分歧，则 $\mathscr{P}|M$；

（2）$H=P_MN_{L/K}(I_M(L))$ 并且由 Artin 映射给出 $I_M/H\cong \mathrm{Gal}(L/K)$.

定理 3 设 L_1/K 和 L_2/K 均为阿贝尔扩张，导子分别为 f_1 和 f_2. 如果 $f_1|M$，$f_2|M$，以 H_i 表示 I_M 中子群 $H_i=P_MN_{L_i/K}\cdot(I_M(L_i))i=1$，2. 则 $H_1\subseteq H_2$ 当且仅当 $L_1\supseteq L_2$.

以上是类域论的主要结果. 特别在定理 2 中取 $f=M=1$，$H=P_1$，对应的域 L 是 K 的最大不分歧（K 的实素除子在 L 中也不分歧）阿贝尔扩张，这即是 K 的希尔伯特类域. 而 Artin 映射给出 $\mathrm{Gal}(L/K)$ 同构于 K 的理想类群 I/P.

现在我们用 Idele 语言叙述类域论. 设 L/K 是数域扩张. 如果 P 和 $\mathscr{P}$ 分别为 L 和 K 的素除子，$P|\mathscr{P}$. 我们以 $N_{P/\mathscr{P}}:L_P\to$

$K_{\mathscr{P}}$ 表示局部域扩张的范映射. 对每个 Idele $x=(x_p)\in J_L$，定义

$$N_{L/K}(x)=(y_{\mathscr{P}})\in J_K,$$

其中，$y_{\mathscr{P}}=\prod_{P|\mathscr{P}}N_{P/\mathscr{P}}(x_p)$. 将 K^x 和 L^x 对角嵌入 J_K 和 J_L 中，则对 K 的每个素除子 $\mathscr{P}$ 和元素 $x\in L^x$，均有 $\prod_{P|\mathscr{P}}N_{P/\mathscr{P}}(x)=N_L/K(x)$(局部范之积为整体范). 所以 $N_{L/K}$把 L^x 映到 K^x 中. 于是得到 Idele 类群的同态 $N_{L/K}:C_L\to C_K$.

定理 1′ 设 L/K 是数域的阿贝尔扩张. 则有群同构

$$J_K/K^xN_{L/K}J_L=C_K/N_{L/K}C_J\cong\mathrm{Gal}(L/K),$$

并且 K 中(有限或无限)素除子 $\mathscr{P}$ 在 L 中分歧当且仅当 $U_{\mathscr{P}}\subseteq K^xN_{L/K}J_L$(这里当 $\mathscr{P}$ 有限时，$U_{\mathscr{P}}$ 是 $K_{\mathscr{P}}$ 中 $\mathscr{P}^-adic$ 单位群，而当 $\mathscr{P}$ 无限时，$U_{\mathscr{P}}=K^x_{\mathscr{P}}$. 并且 $U_{\mathscr{P}}$ 由嵌入 $u_{\mathscr{P}}\mapsto(1,\cdots,u_{\mathscr{P}},\cdots,1)$嵌到 J_K 中).

定理 2′ 若 H 是 C_K 的开子群并且$[C_K:H]$有限，则 K 存在唯一的阿贝尔扩域 L 使得 $N_{L/K}C_L=H$. 换句话说，若 H 是 J_K 的开子群，$K^x\subseteq H$ 并且$[J_K:H]$有限，则 K 存在惟一的阿贝尔扩域 L 使得 $H=K^xN_{L/K}J_L$.

定理 3′ 设 L_1/K 和 L_2/K 均是阿贝尔扩张，则 $L_1\subseteq L_2$ 当且仅当 $K^xN_{L_1/K}J_{L_1}\supseteq K^xN_{L_2/K}J_{L_2}$.

特别地，若 L 是 K 的希尔伯特类域，令 $U=\prod_{\mathscr{P}}U_{\mathscr{P}}$，则 $\mathrm{Gal}(L/K)\cong J_K/K^xU\cong K$ 的理想类群.

有了 Adele 和 Idele 语言之后，韦依在书中使用的进一步工具是 A_K 和 J_K 上的调和分析. 这种方法最早由 Tate 在 1950 年采用，用来证明 $\zeta_K(s)$的函数方程. 韦依则把它系统地同时用于全

部整体域. 我们这里仍以数域情形为例. 首先叙述局部紧交换群上调和分析的一般理论.

任意局部紧群 G 上都有一个测度 μ, 则哈尔(Haar)测度, 这是对 G 作平移不变的测度, 即对 G 的每个可测子集 E 和 $g\in G$, $\mu(gE)=\mu(E)$. 如果 G 是交换群, 则它显然也是右平移不变测度. 每个局部紧群上的哈尔测度本质上是惟一的, 即两个这样的测度相差一个正实数因子. 进而, $\mu(G)$有限当且仅当 G 是紧群, 从而可以有惟一的标准化哈尔测度 μ 使得 $\mu(G)=1$.

设 V 是 $\boldsymbol{C}$ 上拓朴向量空间. 以 $Aut(V)$表示向量空间 V 的自同构群, $\mathrm{Aut}_T(V)$表示向量空间 V 的连续自同构群(当 $\dim V$ 有限时这二者一致). 局部紧群 G 的一个抽象表示是指群同态

$$\rho: G\to \mathrm{Aut}(V),\ g\mapsto \rho_g.$$

而 ρ 叫作连续表示(简称表示), 是指映射

$$G\times V\to V,\ (g, v)\mapsto \rho_g(v)$$

对于 $G\times V$ 的积拓朴是连续的. 这时实际上为 $\rho: G\to \mathrm{Aut}_T(V)$. 由表示 ρ 可将群 G 作用于 V 上, 从而 V 成为 $\boldsymbol{C}[G]$-模. 而每个 $\boldsymbol{C}[G]$-子模 W(即 $\rho(G)W\subseteq W$)叫作 G-不变子空间. 连续表示 ρ 叫作不可约的, 是指 V 只有平凡的 G-不变闭子空间(0)和 V. 如果表示空间 V 是希尔伯特空间(内积表示为$\langle u, v\rangle$), 则 ρ 叫作酉表示是指它保持内积, 即对任意 $u, v\in V$, $\langle u, v\rangle=\langle \rho_g u, \rho_g v\rangle$ (对每个 $g\in G$). 如果 G 是交换群, 则 G 在希尔伯特空间 V 上的不可约酉表示都是一维的(即 $\dim_{\boldsymbol{C}} V=1$), 这种表示叫 G 的(酉)特征, 即为连续同态 $\chi: G\to S^1=\{\alpha\in\boldsymbol{C}: |\alpha|=1\}$. 全体特征自然形成(乘法)交换群, 叫 G 的特征群, 表示成 $\hat{G}$. 可以在 $\hat{G}$ 中引入如下拓朴. 对 G 的每个紧子集 K 和 1 在 S^1 中的邻域 U, 定义

$$W(K, U)=\{\chi\in\hat{G}\mid \chi(k)\subseteq U\}.$$

将全部这种集合作为平凡特征 $\chi\equiv 1$ 的邻域基, 由此决定 $\hat{G}$ 的拓朴叫紧开拓朴. 如果 G 是离散群(即 G 的紧子集均为有限集), 则

上述拓朴即是 $\hat{G}$ 中"逐点收敛"拓朴. 我们有如下的基本结果：

设 G 是局部紧交换群，则 $\hat{G}$ 也为局部紧变换群. 并且

(1) (庞特里雅金对偶)$\alpha: G \to \hat{\hat{G}}$，$\alpha(y)(\chi)=\chi(y)$是拓朴群同构；

(2) G 为离散群当且仅当 $\hat{G}$ 为紧群.

有了局部紧交换群 G 的特征群 $\hat{G}$ 和哈尔测度 dx，可以定义 G 上的付立叶变换. 考虑复值函数空间

$$L'(G) = \{f: G \to \boldsymbol{C} \mid \int_G |f(x)| dx \text{ 收敛}\}.$$

$L'(G)$中函数 $f(x)$的**付立叶变换**是如下定义的函数 $\hat{f}: \hat{G} \to \boldsymbol{C}$，其中对每个 $\chi \in \hat{G}$，

$$\hat{f}(\chi) = \int_G f(y)\overline{\chi}(y) dy.$$

对于 $G=\boldsymbol{R}$(实数加法群)的情形，$\hat{\boldsymbol{R}}=\boldsymbol{R}$(这里 $t \in \hat{R}$ 等同于特征 $\chi: \boldsymbol{R} \to S'$，$\chi(s)=e^{ist}$)，而上面的定义就是通常的付立叶变换. 一般地，我们也有付立叶反变换，即 $\hat{G}$ 上存在(唯一的)哈尔测度 $d\chi$ (叫作 dx 的**对偶**)，使得对每个 $y \in G$，

$$f(y) = \int_{\hat{G}} \hat{f}(\chi)\chi(y) d\chi.$$

以上是局部紧交换群上调和分析的基本结果. 现在把这些结果具体应用于 Adele 群 $\boldsymbol{A}_K$ 和 Idele 群 J_K 上. 我们可以在较一般的情形下进行，结果非常简洁.

设 I 是一个无限集合，I_∞是 I 的一个有限子集合. 对每个 $V \in I$，我们有一个局部紧群 G_V. 对每个 $V \in I - I_\infty$，则 G_V 有一个紧开子群 H_V. 代数上我们考虑直积 $\prod_{V \in I} G_V$ 的一个子群

$$G = \prod{}' G_V = \{(x_V) \in \prod_{V \in I} G_V \mid x_V \in G_V \text{，对几乎所有 } V \text{，} x_V \in$$

$H_V\}$，

G 叫作$\{G_V\}$对$\{H_V\}$的**限制直积**. 考虑 G 的子群 $\prod_V N_V$，其中 N_V 是 1 在 G_V 中的邻域，并且对几乎所有 V，$N_V=H_V$. 将所有这种子群看成 1 在 G 中的邻域基，则 G 成为拓朴群，并且是局部紧群. 每个 G_V 通过投射均看成 G 的闭子群. 不难看出，数域 K 上的 Adele 环和 Idele 群都是这种限制直积的特例. 下面是限制直积 $G=\prod{}'G_V$ 上调和分析的基本结果.

(1) 关于特征群 $\hat{G}$. 若 $\chi:G\to C^x$ 为连续同态，则对几乎所有 V，$\chi_V(H_V)=1$，其中 χ_V 表示 χ 在 G_V 上的限制. 记 $H_V^*=\{\chi\in\hat{G}_V|\chi(H_V)=1\}$，这是 $\hat{G}_V$ 的紧开子群. 于是 $\chi=\prod_V\chi_V$，$\chi_V\in\hat{G}_V$，并且对几乎所有 V，$\chi_V\in H_V^*$. 由此可证 $\hat{G}\cong\prod{}'\hat{G}_V$，右边是$\{\hat{G}_V\}$对$\{H_V^*\}$的限制直积(注意 H_V^* 拓朴同构于$(G_V/H_V)\hat{}$).

(2) G 上的哈尔测度. 若对每个 V，$\mathrm{d}g_v$ 是 G_V 的哈尔测度，并且对几乎所有 V，$\int_{H_V}\mathrm{d}g_v=1$. 则 G 上有唯一的哈尔测度 $\mathrm{d}g$，使得

$$\int_G f(g)\mathrm{d}g=\prod_V(\int_{G_V}f_V(g_V)dg_V(\text{对每个 } f\in L'(G)),$$

其中 f_V 表示 f 在 G_V 上的限制. 这个哈尔测度 $\mathrm{d}g$ 表示成 $\prod_V\mathrm{d}g_V$.

(3)G 上的付立叶分析. 若 G_V 均为局部紧交换群. 对于 $g=(g_V)\in G$，则 $\hat{f}(g)=\prod_V\hat{f}_V(g_V)$. 进而，以 $\mathrm{d}\chi_V$ 表示 $\mathrm{d}g_V$ 的对偶测度，则对几乎所有 V，$\int_{H_V^*}\mathrm{d}\chi_V=1$. 于是可以定义 $\hat{G}=\prod{}'\hat{G}_V$ 上的测度 $\mathrm{d}\chi=\prod_V\mathrm{d}\chi_V$. 这时，设 $f=\prod f_V$，并且对几乎所有

V, $f_V(H_V)=1$. 则有付立叶反变换 $f(g)=\int_{\hat{G}}\hat{f}(\chi)\chi(g)\mathrm{d}\chi$.

以上是限制直积 $G=\prod{}'G_V$ 上调和分析的基本结果. 简言之, G 的特征群、哈尔测度和付立叶变换都是把诸 G_V 的相应信息以简洁的形式收集在一起. 具体到 Adele 环 $\boldsymbol{A}_K=\prod_{\mathscr{P}}{}' K_{\mathscr{P}}$ 和 Idele 群 $J_K=\prod_{\mathscr{P}}{}' K_{\mathscr{P}}^{\times}$ 上, Tate 计算了每个局部域加法群 $K_{\mathscr{P}}$ 和乘法群 $K_{\mathscr{P}}^{\times}$ 的特征群、哈尔测度和付立叶变换与反变换公式. 把这些信息收集到一起, 便得到 $\boldsymbol{A}_K$ 和 J_K 的特征群、哈尔测度和付立叶变换与反变换公式, 其中 K 是任意代数数域. 作为应用, Tate 用这种工具计算出 K 上整体 zeta 函数

$$Z(f,\chi)=\int_{J_K}f(x)\chi(x)\mathrm{d}^*x$$

的解析延拓和函数方程, 其中 χ 是 J_K 的任意复特征, $\chi(K^x)=1$ (这叫作 idele 类的特征), 而 f 是 $\boldsymbol{A}_K$ 上性状良好的复函数, $\mathrm{d}^x x$ 是 J_K 上的哈尔测度. 这种整体 zeta 函数是前人所研究的一大批 zeta 函数和 L-函数的推广.

让我们再回忆一下历史. 黎曼对 $\boldsymbol{Q}$ 上的 zeta 函数 $\zeta(s)=\sum\limits_{n\geqslant 1}n^{-s}$ 给出函数方程和解析延拓, 他的方法是把级数 $\sum\limits_{n\geqslant 1}n^{-s}$ 表成积分形式, 公式中出现伽马函数 $\Gamma(s/2)$ 因子, 而被积函数是 zeta 函数 $\theta(z)=\sum\limits_{n\in\boldsymbol{Z}}e^{2\pi in^2z}$. 利用实数域上的 Poisson 求和公式得到变换公式 $\sum\limits_{n\in\boldsymbol{Z}}e^{-\pi n^2t^2}=t^{-1}\sum\limits_{n\in\boldsymbol{Z}}e^{-\pi n^2/t^2}\ (t>0)$, 给出函数方程

$$\zeta(s)=\zeta(1-s),\ \zeta(s)=\pi^{-s/2}\Gamma(s/2)\zeta(s).$$

采用类似的方法, 狄里赫利得到 L-函数 $L(s,\chi)=\sum\limits_{n\geqslant 1}\chi(n)n^{-s}$ 的函数方程和解析延拓, 其中 χ 是狄里赫利特征. 把它们加以推广, Hecke 在 20 世纪 20 年代考虑任意数域 K 上的 L-函数 $L(s,\chi)=\sum\chi(A)NA^{-s}$, 其中 A 过 O_K 的所有非零整理想, 而 χ 是

K 的理想类群(有限交换群)的特征(当 $\chi\equiv1$ 时,这就是狄德金 zeta 函数 $\zeta_K(s)=\sum N(A)^{-s}$). Hecke 甚至作了更大的推广,用现在的语言表述,他取 χ 为 idele 类群 $C_K=J_K/K^\times$ 的特征. Hecke 使用 poisson 公式的高维模拟,得到高维 theta 函数的变换公式,推出 Hecke L-函数的函数方程 $L(s,\chi)=W(\chi)L(1-s,\hat{\chi})$. 但是 Hecke 对 $W(\chi)$(叫作根数,root number)没有给出满意的描述. 1950 年前后,Tate 在导师 E. Artin 的建议下,采用 A_K 上的付立叶分析工具(包括 A_K 上的 Poisson 公式),利用 $L(s,\chi)$的无穷乘积展开,将 $L(s,\chi)$分解成局部因子乘积. 在每个局部域上,局部因子有简单的函数方程. 将它们拼在一起,便得到 Hecke L-函数的整体函数方程,而根数 $W(\chi)$是局部函数方程中出现的局部常数的乘积,这就给出根数 $W(\chi)$一个更为清晰的刻划. 韦依在《基础数论》一书中对于数域和函数域统一地作了类似的工作. 对于函数域情形,韦依对 $L(s,\chi)$的零点证明了黎曼猜想的模拟,并且用 A_K 的 Poisson 公式得到黎曼-洛赫定理. 详细计算可见所列参考书.

Adele 环和 Idele 群上的分析工具不仅给出处理经典数论问题的新的统一方法(除了函数方程之外,还可导出类数解析公式以及各种 L-函数的解析特性和密度定理等),而且有许多现代应用. 其中一个最重要的应用是对任意约化(reductive)的矩阵群 G 研究它的"Adele 化"$G(A_K)$. 典型的例子是对每个数域(或函数域)K, $G=GL_n(K)$(K 上的 n 次一般线性群). 它的"Adele 化"为 A_K 上 n 阶可逆方阵组成的乘法群 $GL_n(A_K)$. 当 $n\geqslant2$ 时这是非交换的局部紧群(而 $GL_1(A_K)=J_K$). 利用对角嵌入可把 $GL_n(K)$看成 $GL_n(A_K)$的子群,而 $GL_n(K)/GL_n(A_K)$是重要的齐性空间. 以 Z_n 表示群 $GL_n(A_K)$的中心,则齐性空间 $X_K^{(n)}=GL_n(K)Z_n/GL_n(A_K)$对于其上的哈尔测度 μ, $\mu(X_K^{(n)})$是有限的. 以 $L^2(X_K^{(n)})$表示 $X_K^{(n)}$ 上平方可积复值函数组成的希尔伯特空间,群 $GL_n(A_K)$

自然地(右)作用在此空间上给出此群的表示. 根据下节要介绍的朗兰兹猜想，这些表示和空间 $L^2(X_K^{(n)})$上的调和分析是理解数域(或函数域)K 所有非阿贝尔扩张的关键.

参 考 文 献

[1] F. Bruhat, Lectures on Lie Groups and Representations of Locally Compact Groups, Bombay: Tata Institute of Fundamental Research, 1968.

[2] W.C. Winnie Li, Number Theory with Applications, World Scientific, 1996.

[3] D. Ramakrishnan, R. J. Valenza, Fourier Analysis on Number Fields, GTM 186, Springer-Verlag, 1999.

[4] J. Tate, Fourier Analysis in Number Fields and Hecke's Zeta Functions, Thesis, Princeton University, 1950.

[5] A. Weil, Basic Number Theory(第三版英文版), Springer-Verlag, 1974.

[6] J.-P. Serre, La vie et l'oeuvre d'Andre Weil, L'enseignement Mathematique, 45(1999), 5～16.

§4.2 朗兰兹猜想(1967年)

1983年，美国科学院为美国总统的科学技术委员会提交了一份《理论数学的进步》年度报告，在综述部分有这样两段话：

“费利克斯·克莱因(Felix Klein)1872年在他的爱尔朗格纲领(Erlanger program)所透彻阐释的几何学中群对称的统一作用导致一个世纪的数学进步. 与爱尔朗格纲领相媲美的后继应当是朗兰兹纲领，它是用李群的无穷维表示来阐释数论.

朗兰兹猜想是用群$GL(n)$的无穷维不可约表示来刻划n次数域的性状，而他的更宏伟的猜想提出一系列使人眼花撩乱的问题. 这些问题的解决会使我们对于表示论、数论和代数几何有更好的理解. 这方面已经取得了引人注目的进展，但还有更多的事情留给了未来.”

朗兰兹纲领是从1967年起由美国数学家朗兰兹(Langlands，1940—)以一系列猜想的形式提出的. 这些猜想的本质是试图发现复的和p-adic李群的无穷维表示理论、调和分析、代数几何与数论之间的深层联系. 这些猜想对于纯粹数学的一个广阔的领域表达出一种普遍而富有哲理的观点，被人们称之为朗兰兹哲学.

尽管朗兰兹猜想高深莫测，但仍旧根源于古典数论的基本问题，即三千年前人类就开始研究的代数方程有理解和整数解问题. 对于整系数的多项式方程$f(x_1, \cdots, x_n)=0$，人类很早就知道用同余方程$f(x_1, \cdots, x_n)\equiv 0 \pmod m$($m$为任意正整数)来考查方程$f=0$是否有整数解的可能性. 到了19世纪初期，高斯把

二次方程 $x^2-d\equiv 0(\bmod p)$的可解性(即 x^2-d 在有限域 $\boldsymbol{F}_P$ 中的分解特性)和素数 p 在二次域 $\boldsymbol{Q}(\sqrt{d})$中的分解特性联系起来. 19 世纪中期，伽罗华研究整系数代数方程 $f(x)=0$ 的根式可解性，他考虑 $f(x)$的根之间可能的置换，从而产生了(置换)群的概念. 如果 $f(x)$是整系数 n 次不可约多项式，K 是 $f(x)$的分裂域，即 $K=\boldsymbol{Q}(\alpha_1, \cdots, \alpha_n)$，其中 $\alpha_1, \cdots, \alpha_n$ 为 $f(x)$在 $\boldsymbol{C}$ 中的 n 个根，则 $K/\boldsymbol{Q}$ 是伽罗华扩张. 伽罗华群 $G=\mathrm{Gal}(K/\boldsymbol{Q})$就是 $f(x)$的诸根 $\alpha_1, \cdots, \alpha_n$ 的置换群，也把 G 叫作是多项式 $f(x)$的伽罗华群. 到了 19 世纪后期，代数数论有了长足的进步之后，人们认识到高斯对二次多项式的研究具有普遍意义，即除了有限个例外情形，对每个素数 p，p 在 K 中不分歧，并且若理想 pO_K 在 O_K 中分解成 g 个素理想的乘积，则 $f(x)$模 p 分解成 g 个次数均为 f 的 $\boldsymbol{F}_p$ 上不可约多项式的乘积，其中 $fg=[K:\boldsymbol{Q}]=|G|=degf(x)$，而 f 是群 G 中关于 p 的 Frobenius 元素的阶.

如果 G 是(有限)阿贝尔群，即 K 是阿贝尔数域，Kronecker—Weber 定理告诉我们，K 必是某个分圆域 $\boldsymbol{Q}(\zeta_m)$的子域，从而 G 同构于$(\mathbf{Z}/m\mathbf{Z})^x$ 的某个商群. 当素数 p 与 m 互素时，关于 p 的 Frobenius 元素的阶即是 p 在这个商群中的阶，从而只与 p 模 m 的同余类有关. 由此可以得出 $f(x)$模 p 分解的十分清楚的法则，并且分解模式(即 $f(x)$在 $\boldsymbol{F}_p$ 上分解不可约多项式个数 g 和它们的次数 f)只与 p 的模 m 同余类有关. 人们把 p 在 K 中的这种分解规律(也就是 $f(x)(\bmod p)$分解法则)叫作阿贝尔域 K 的互反律，它是二次域中 p 的分解规律(即古典的二次互反律)的推广.

希尔伯特第 9 问题是:对于任意代数数域 K 应当有什么样的互反律可看作是古典二次互反律的推广? 用更朴素的语言，这个问题相当于:对于每个整系数不可约多项式 $f(x)$，如何用 $f(x)$ 的自身特性来描述 $f(x)$在每个有限域 $\boldsymbol{F}_p$ 上的分解规律? 代数数域告诉我们，分解规律依赖于 f 的伽罗华群 $G=\mathrm{Gal}(K/\boldsymbol{Q})$，其中 K 是 f 的分裂域. 如果 G 不是阿贝尔群，则 $f(x)(\bmod p)$的分解

规律如何由 $f(x)$（或者由 K）来刻划，至今是一个未解决的问题.

更一般地，设 K 为任意数域，$f(x)$ 是 $O_K[x]$ 中的不可约多项式，L 是 $f(x)$ 在 K 上的分裂域. 则除了有限个情形之外，O_K 中每个素理想 P 在 L 中不分歧. 并且若 PO_L 分解成 O_L 中 g 个不同素理想之积，则 $f(x)$ 在有限域 O_K/P 上分解成 g 个次数均为 f 的不可约多项式之积，其中 $fg=[L:K]$，并且 f 是 Frobenius 元素 $\left(\frac{L/K}{P}\right)$ 在伽罗华群 $G=Gal(L/K)$ 中的阶. 问题为：如何用基域 K 的特性来描述这些 Frobenius 元素？当 L/K 是阿贝尔扩张时，类域论中的 Artin 互反律给出这问题的一个解答. 对于 L/K 是非阿贝尔扩张的情形，至今没有满意的解答. 类域论的主要功能是用数域 K 的自身特性（开始时用 K 的广义类群，后来用 idele 群 J_K）刻划 K 的所有阿贝尔扩张. 现在的问题是如何用 K 的自身特性刻划 K 的所有伽罗华扩张. K 的所有阿贝尔扩张的合成是 K 的最大阿贝尔扩张 K^{ab}，类域论本质上刻划了无限伽罗华群 $\mathrm{Gal}(K^{ab}/K)$. 而 K 的所有伽罗华扩张的合成即是 K（在 C 中）的代数闭包 K^{ac}，我们需要研究无限伽罗华群 $G=\mathrm{Gal}(K^{ac}/K)$. 对于 K 的每个有限伽罗华扩域 L，$G_L=\mathrm{Gal}(K^{ac}/L)$ 是 G 的子群并且 $[G:G_L]$ 有限. 把所有这种 G_L 看作是 G 中幺元素（恒等自同构）的邻域基，由此使 G 为拓朴群（Krull 拓朴）. 根据无限伽罗华扩张的基本定理，G 的所有闭子群 H 和 K^{ac}/K 的所有中间域 L 是一一对应的，并且 L/K 为有限扩张当且仅当 H 是 G 的开子群，L/K 是伽罗华扩张当且仅当 H 是 G 的正规子群. 所以，要决定 K 的全部有限伽罗华扩域，本质上归结于一个问题，即用 K 的自身特性来刻划伽罗华（拓朴）群 $G=\mathrm{Gal}(K^{ac}/K)$ 的结构. 在群表示理论出现之后，人们发现用 G 的表示来研究群 G 结构是非常有效的手段. 当 L/K 为有限阿贝尔扩张时，有限阿贝尔群 Gal (L/K) 同构于它的特征群（特征即是此群的不可约表示），问题是简单的. 对于任意有限伽罗华扩张 L/K，群 $\mathrm{Gal}(L/K)$ 非交换时，

则存在高维不可约表示，使问题复杂化. 这些有限非阿贝尔群 Gal(L/K)的表示以及无限拓朴群 Gal(K^{ac}/K)的表示都叫作伽罗华表示. 研究这些伽罗华表示成为现代数论的中心问题之一.

早在 1923 年，E. Artin 就在数域任意伽罗华扩张 L/K 的研究中引进群表示方法. 设 L/K 是有限伽罗华扩张，$\rho: G\to GL_n(C)$是伽罗华群 $G=\mathrm{Gal}(L/K)$的一个 n 维复表示. 除了有限个例外情形，O_K 中素理想 p 在 L 中不分歧，Frobenius 元素 $F_P=\left(\frac{L/K}{P}\right)$是群 G 的一个共轭元素类. 根据有限群表示理论，这个共轭类由它们的共同特征多项式 $\det(I_n-\rho(F_P)N(P)^{-S})$所完全决定的. Artin 引进伽罗华扩张 L/K 关于表示 ρ 的 L-函数.

$$L(s, \rho)=\prod_P \det(I_n-\rho(F_P)N(P)^{-s})^{-1}$$

(对于例外情形的 P 需要适当设定乘积中的 p-因子). 如果 L/K 是阿贝尔扩张而 ρ 是一维表示，则 $L(s, \sigma)$本质上为 Hecke L-函数，$K=Q$ 时这就是狄里赫利 L-函数. Artin 用这种方法于 1925 年前后给出阿贝尔扩张的互反律. 当 G 是非阿贝尔群时，Artin 证明了 $L(s, \sigma)$的一系列解析性质，但是他不能发现狄里赫利特征和狄里赫利 L-函数的高维模拟，不知道 G 的高维表示如何用 K 的自身特性去体现. 有趣的是，就在同一时间(1927 年)，和 Artin 在同一个学校(汉堡大学)工作的 Hecke 研究了模形式的 L-函数. 到了 1951 年，韦依用 idele 语言叙述类域论，数域的阿贝尔扩张 L/K 的伽罗华群 Gal(L/K)同构于 idele 类群 C_K 的某个商群. 为了研究任意伽罗华扩张 L/K，他构作了一个新的群(现在叫作韦依群)由此得到一种新型的 L-函数，而 Artin 的非阿贝尔 L-函数和 Hecke 关于模形式的 L-函数都是它的特例，就像韦依所说的，"实现了 Artin 和 Hecke 的联姻". Artin L-函数和 Hecke 关于模形式 L-函数之间的明确联系是朗兰兹提出的. 这时，Hecke 模形式已经被看成是群 $GL_2(A_Q)$的自守形式，而模形式的 L-函数是此群的自守表示的 L-函数. 朗兰兹认为群

$GL_2(A_K)$的自守表示是狄里赫利特征的适宜的推广，这些自守表示的 L-函数是狄里赫利 L-函数的推广. 1967 年朗兰兹提出了他的第一个猜想：

朗兰兹第一猜想 设 L/K 是数域的伽罗华扩张，$G=\mathrm{Gal}(L/K)$，$\rho:G\to GL_n(\boldsymbol{C})$是 G 的 n 次不可约复表示，$L(s,\rho)$是 Artin L-函数. 则存在群 $GLn(\boldsymbol{A}_K)$的尖点型自守表示 π_ρ，使得这个自守表示 π_ρ 的 L-函数 $L(s,\pi_\rho)$等于 $L(s,\rho)$.

确切地叙述自守表示需要很多数学概念和语言，我们这里只打算作描述性地解释，并且沿着历史发展的途径，说明如何把模形式看成是自守形式并与无穷维的自守表示发生联系. 据 S. Gelbart 考证，这件事是 1952 年由前苏联数学家 I. M. Gelfand 和 S. Fomin 在题为《常负曲率流形上的测地流》一文中首次给出的.

设 $f(z)$是群 $SL_2(\boldsymbol{Z})$的权 k 模形式. 则 $f(z)$是上半平面 $\boldsymbol{H}$ 中(包括无穷远点 $i\infty$)的全纯函数并且对$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\in SL_2(\boldsymbol{Z})$，

$$f\left(\frac{az+b}{cz+d}\right)=(cz+d)^{-k}f(z) \qquad (z\in \boldsymbol{H}). \tag{1}$$

整个群 $SL_2(\boldsymbol{R})$作用于上半平面 $\boldsymbol{H}$，$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$把 z 映成$\frac{az+b}{cz+d}$，这个作用是可迁的. $\boldsymbol{H}$ 中点 i 的固定子群为 $SO_2(\boldsymbol{R})=\left\{\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \middle| 0\leqslant\theta<2\pi\right\}$. 于是

$$\boldsymbol{H}=SL_2(\boldsymbol{R})/SO_2(\boldsymbol{R})$$

对每个关于群 $SL_2(\boldsymbol{Z})$的权 k 模形式 $f(z)$，定义 $SL_2(\boldsymbol{R})$上的复值函数 ϕ_f 为：

$$\phi_f(g)=(ci+d)^{-k}f\left(\frac{ai+b}{ci+d}\right) \quad \left(g=\begin{pmatrix} ab \\ cd \end{pmatrix}\in SL_2(\boldsymbol{R})\right). \tag{2}$$

容易验证，若 $r(\theta)=\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\in SO_2(\boldsymbol{R})$，则($\phi_f$ 简记为 ϕ)

$$\phi(gr(\theta))=e^{-ik\theta}\phi(g). \tag{3}$$

公式(1)转化成

(i) 若 $r\in SL_2(\boldsymbol{Z})$，则 $\phi(rg)=\phi(g)$.

f 的全纯性质相当于：对于 $SL_2(\boldsymbol{R})$上的拉普拉斯算子 Δ(这是二阶微分算子)

(ii) $\triangle\phi=-\frac{1}{4}k(k-2)\phi$.

局部紧群 $SL_2(\boldsymbol{R})$上的哈尔测度自然诱导出商空间 $SL_2(\boldsymbol{Z})/SL_2(\boldsymbol{R})$上的哈尔测度. 以 L^2 表示这个商空间上对这个哈尔测度平方可积函数全体，由(i)知 $\phi\in L^2$. 如果 f 是尖点模形式，即 $f(i\infty)=0$，则可转化成 ϕ 的"尖点条件"，其中包括

(iii) ϕ 的"第 0 个 Fourier 系数"

$$\int_{Z/R}\phi\left(\begin{pmatrix}1 & x\\ 0 & 1\end{pmatrix}g\right)\mathrm{d}x$$

等于 0. 这样的 ϕ 形成 L^2 的一个子空间 L_0^2. 反之，如果 $SL_2(\boldsymbol{R})$上的函数 $\phi(g)$满足公式(3)和上述三个条件，则必存在关于 $SL_2(\boldsymbol{Z})$的权 k 尖点模形式 f，使得 $\phi=\phi_f$. 如果只把条件(ii)减弱成"ϕ 是算子 Δ 的本征函数"(即本征值不必为$-\frac{1}{4}k(k-2)$)，我们便把满足上述诸条件的函数叫作 $SL_2(\boldsymbol{R})$上的**自守尖点形式**. 除了 $\phi_f(f\in S_k(\Gamma)$，$\Gamma=SL_2(\boldsymbol{Z}))$之外，这种扩大的定义确实包含新的自守尖点形式，这就是 H. Maass 在 1949 年研究实二次域狄德金 zeta 函数时发现的实解析自守尖点形式(wave forms).

下一步要解释 $SL_2(\boldsymbol{R})$上的自守形式和群 $SL_2(\boldsymbol{R})$的表示之间的关系. 设 ρ 是群 $G=SL_2(\boldsymbol{R})$在空间 L_0^2 上的右正则表示，即对于 $\phi\in L_0^2$ 和 g，$h\in SL_2(\boldsymbol{R})$，

$$\rho(g)\phi(h)=\phi(hg).$$

这实际上是 G 的酉表示，并且对于足够好的 ϕ，算子 Δ 与表示 ρ 可交换：$\Delta\rho\phi=\rho\Delta\phi$. 根据表示理论，可以用算子 Δ 把表示 ρ 分解

成一些子表示的直和.

例如对每个尖点模形式 $f\in S_k(\Gamma)$，以 L_f 表示由 ϕ_f 和它的 $SL_2(\boldsymbol{R})$右平移作用所得函数全体张成的 L_0^2 中的子空间. 如果 f 是 Hecke 模形式，则 L_f 是 L_0^2 对于表示 ρ 的不可约子空间. 用这种方式得到 $SL_2(\boldsymbol{R})$的一族不可约表示，叫作权 k 的离散表示系列. 更一般地，自守尖点形式按这种方式对应着 ρ 在 L_0^2 中的一个子表示，它是不可约酉表示，称之为**自守尖点表示**. 换句话说，使用自守形式理论可以知道右正则表示 ρ 是如何分解成不可约酉表示的直和.

事情还不止于此，实数域 $\boldsymbol{R}$ 只是有理数域 $\boldsymbol{Q}$ 的一个局部分支 $\boldsymbol{Q}_\infty$. 我们还需利用局部-整体原则把一个局部域 $\boldsymbol{R}=\boldsymbol{Q}_\infty$ 扩大到 Adele 环 $\boldsymbol{A}=\boldsymbol{A}_{\boldsymbol{Q}}=\prod_p{}'\boldsymbol{Q}_p$(包含 $p=\infty$)上. 对每个素数 p，考虑

$$G_p=GL_2(\boldsymbol{Q}_p),\ K_p=GL_2(\boldsymbol{Z}_p),$$

则 K_p 是 G_p 的紧子群. 再记 $G_\infty=GL_2(\boldsymbol{R})$，便可考虑$\{G_p\}$对$\{K_p\}$的限制直积(这是局部紧群)

$$G_A=GL_2(\boldsymbol{A})=\prod_{P\leqslant\infty}{}' G_p.$$

采用对角嵌入可把 $G_Q=GL_2(\boldsymbol{Q})$看成是 G_A 的离散子群.

对于每个尖点模形式 $f(z)\in S_k(\Gamma)$，类似于前面公式(2)的方式可把 $f(z)$提升成 G_A 上的一个函数 ϕ_f，并且对每个 $r\in G_Q$，$\phi(rg)=\phi(g)$. 对 G_A 的中心 $Z_A=\{aI_2|a\in J=J_Q\}$中元素 z，亦有 $\phi(zg)=\phi(g)$. 于是 $\phi=\phi_f$ 可看成是 $Z_AG_Q\backslash G_A$ 上的函数. 于是又考虑 $G_A=GL_2(\boldsymbol{A})$在空间 $L_0^2=L_0^2(Z_AG_Q\backslash G_A)$上的右正则表示 ρ，对每个 $f\in S_k(\Gamma)$，ϕ_f 和它的所有右平移生成 L_0^2 的不变子空间，并给出 ρ 的不可约子表示. 当 $n=1$ 时，$G_A=GL_1(\boldsymbol{A})=J$，$Z_AG_Q\backslash G_A=\boldsymbol{Q}^x\backslash J=C_Q$($\boldsymbol{Q}$ 的 idele 类群)，而 G_A 的不可约表示就是 idele 类特征，这种特征都可表示成局部特征之积 $\chi=\prod_p\chi_p$. 类似地对 $n=2$ 的情形，也可证明 G_A 的每个不可约酉表示 π 都可表成张量

积$\otimes_p \pi_p$，其中 π_p 是群 G_p 的不可约酉表示．如果 f 是 Hecke 权 k 模形式，由 ϕ_f 按上述方式给出不可约表示 $\pi=\otimes\pi_p$．对每个 Hecke 算子 $T(p)$，$T(p)f=a_p f$，则局部分量 π_p 可由 a_p 和 k 完全描述．

上面的理论还可作进一步的推广，即将有理数域 $\boldsymbol{Q}$ 改成任意代数数域（或函数域）K，再将群 GL_2 改成 GL_n（甚至任意连通约化代数群 G 和它的离散子群 Γ，使得 Γ/G 具有有限测度）．对 K 的每个素除子 $\mathscr{P}$，记 $G_{\mathscr{P}}=GL_n(K_{\mathscr{P}})$．若 $\mathscr{P}$ 是有限的（即 $\mathscr{P}$ 为 O_K 的素理想），则 $M_{\mathscr{P}}=GL_n(O_{\mathscr{P}})$ 是 $G_{\mathscr{P}}$ 的紧子群，其中 $O_{\mathscr{P}}$ 是局部域 $K_{\mathscr{P}}$ 的整数环．以 $G_A=G(A_K)=\prod\limits_{\mathscr{P}}{}' G_{\mathscr{P}}$ 表示 $\{G_{\mathscr{P}}\}$ 对于 $\{M_{\mathscr{P}}\}$ 的限制直积．由对角嵌入可把 $G_K=GL_n(K)$ 看成 G_A 的离散子群．以 L_0^2 表示 $G_K\backslash G_A$ 上平方可积并且满足某种尖点条件的复值函数所构成的希尔伯特空间，ρ 是 G_A 在空间 L_0^2 中的右正则表示，则把 ρ 中出现的每个不可约酉表示都叫作 G_A 的**自守尖点表示**，每个这样的表示均可表成 $\pi=\otimes\pi_{\mathscr{P}}$，其中 $\pi_{\mathscr{P}}$ 是 $G_{\mathscr{P}}$ 的不可约酉表示．对几乎所有 $\pi_{\mathscr{P}}$ 可构作一个局部 L-函数 $L(s,\pi_{\mathscr{P}})=P(N_{\mathscr{P}}^{-s})^{-1}$，其中 $P(z)$ 是常数项为 1 的 n 次多项式．对有限个例外的 $\mathscr{P}$ 也可适当定义 $L(s,\pi_{\mathscr{P}})$．然后可定义自守尖点表示 π 的 L-函数

$$L(s,\pi)=\prod_{\mathscr{P}} L(s,\pi_{\mathscr{P}}).$$

右边的无穷乘积在某个右半平面上是收敛的，而 Jacquet 和朗兰兹于 1970 年对于 $n=2$ 情形，Godement 和 Jacquet 于 1972 年对于一般情形证明了 $L(s,\pi)$ 可以解析延拓成整个复平面上的全纯函数（惟一的例外是 $n=1$ 和 $\pi=1$，这时 $L(s,\pi)$ 有一个极点），并且有函数方程

$$L(s,\pi)=\varepsilon(s,\pi)L(1-s,\tilde{\pi}),$$

其中 $\tilde{\pi}$ 是与 π 有关的另一个自守表示，而因子 $\varepsilon(s,\pi)$ 可表成局部

因子的乘积：$\varepsilon(s, \pi) = \prod_{\mathscr{P}} \varepsilon(s, \pi_{\mathscr{P}})$.

对于 $n=1$ 的情形，以上结果就是我们上节介绍的 Tate 于 1950 年的工作. 对于 $n=2$，上述结果是 Hecke 关于模形式 L 函数结果的推广.

有了以上这些准备，让我们再回过头来看一下朗兰兹第一猜想的意义. 这个猜想是说：对于任意数域伽罗华扩张 L/K，它对于 $\mathrm{Gal}(L/K)$的 n 维不可约复表示 ρ 的 Artin L-函数 $L(s, \rho)$一定是 $GL_n(A_K)$对某个自守表示 π 的 L-函数 $L(s, \pi)$. 当 $n=1$ 并且 L/K 是阿贝尔扩张时，这个结果就是 Artin 的互反律，从而这个猜想相当于类域论的高维推广. 另一方面，这个猜想是用 K 的自身特性（即 $GL_n(A_K)$的自守表示）来刻划所有伽罗华扩张 L/K，可以看成是任意数域 K 的一般互反律，即这个猜想是希尔伯特第 9 问题的一种解答（尽管此猜想对于 $n\geqslant 2$ 情形至今仍未证明）. 最后，Artin 猜想在不可约表示 ρ 的次数 $n\geqslant 2$ 时，Artin L-函数 $L(s, \rho)$均是全纯函数. 上述朗兰兹猜想可以推出 Artin 这个猜想，因为已经证明了对应自守表示的 L-函数都是全纯的.

1970 年以后，朗兰兹考虑其他复约化李群（如酉群 $SO_{2n+1}(\boldsymbol{C})$和辛群 $SP_{2n}(\boldsymbol{C})$等），这些群上的自守表示 π 和数域伽罗华扩张 L/K 相联系，并且猜想这些自守表示的 L-函数可以解析延拓成复平面上亚纯函数并且有函数方程（朗兰兹第 2 猜想），而且具有某种函子性质（朗兰兹第 3 猜想）. 进而，这些猜想也都可以在函数域上加以考虑.

综上所述，朗兰兹纲领是对现在数学诸多领域一种统一性的看法和普遍性的观点，由一系列规模宏大的猜想所组成，其中有些猜想甚至还没有形成很明确的数学语言. 解析方法在欧拉、黎曼、狄里赫利和狄德金年代成功地引入数论研究中以后，对于数论的重要研究对象 X（如数域的伽罗华扩张，数域上代数簇，椭

圆曲线或阿贝尔簇有理点群结构等)，人们试图寻找一个解析函数 $L(X, s)$，使得这个解析函数的各种解析特性(解析延拓，无穷乘积展开，函数方程，零点和极点特性，在某些特殊点的取值等)能充分反映该数论对象 X 的算术、代数甚至几何或组合特性(例如 $L(X, s)$的无穷乘积体现 X 上的局部-整体原则，函数方程体现 X 的某种对称性……). 朗兰兹纲领哲理性地告诉我们这样的解析函数来源于何方:应当存在与 X 有关的适当的环 $R(X)$，使得所求的解析函数为一般线性群 $GL_n(R(X))$或者 $R(X)$上更一般代数群的某种自守表示 π 的 L-函数 $L(s, \pi)$.

迄今为止，对于高维情形($n\geqslant 2$)只对少数例子证明了朗兰兹猜想，多数工作是 $n=2$ 情况，关于 $n\geqslant 3$ 情形的研究最近才有较大的开展. 对于数域伽罗华扩张，目前只对某些非交换伽罗华群(如对称群 S_4 和交错群 A_4)的二维不可约表示，证明了朗兰兹第1猜想. 对于数域 K 上的代数簇，只对椭圆曲线情形提出了明确的谷山-志村-韦依猜想，把椭圆曲线的 L-函数与某种 Hecke 权 2 模形式的 L-函数联系起来. 1994 年怀尔斯对于 $\boldsymbol{Q}$ 上半稳定的椭圆曲线证明了谷山-志村-韦依猜想，并由此推出费马猜想(见§4.6). 另一方面，对于函数域情形却有较大的突破，前苏联数学家 Drinfeld 在 20 世纪 70 年代中期对于朗兰兹第 1 猜想证明了函数域的 $n=2$ 情形(见§4.5). 这项工作和 Drinfeld 在量子群和数学物理等方面的工作得到 1990 年 Fields 数学奖，而朗兰兹和怀尔斯一起获得 1996 年沃尔夫奖，可以预料，朗兰兹纲领将带动 21 世纪理论数学的发展和进步.

参 考 文 献

[1] E. Artin, Zur Theorie der L-Reihen mit allgemeineren Gruppen-Charakteren, Abh. Math. Sem. Univ. Hamburg 8 (1930), 292～306.

[2] S. Gelbart, Automorphic forms on Adele Groups, Princeton Univ. Press, 1975.

[3] ——, An elementary introduction to the Langlands program, Bull. Amer. Math. Soc. 10(1984), 177～219.

[4] I. Gelfand, M. Graev and I. Piatetski-Shaplro, Representations Theory and Automorphic Functions, Academic Press, 1990.

[5] R. Godement, H. Jacquet, Zeta Functions of simple algebras, Lecture Notes in Math. 260, 1972.

[6] E. Hecke, Eine neur Art von Zetafunktionen und ihre Beziehung zur Verteilung der Drimzahlen, M. Z. 1(1918), 357～376, 6(1920), 11～51.

[7] H. Iwaniec, Introduction to the Spectral Theory of Automorphic Forms, Revista Math. Iberoamericano, Madrid, 1995.

[8] H. Jacquet, Automorphic Forms on GL(2), part Ⅱ, Lecture Notes in Math. 278, 1972.

[9] H. Jacquet and R. Langlands, Automorphic Forms on GL(2), Lecture Notes in Math. 114, 1970.

[10] S. Lang, SL(2, **R**), Springer-Verlag, 1985.

[11] R. P. Langlands, L-functions and automorphic representations, Proc. Internat. Congress Math., Helsinki, 1978.

[12] D. Bump, Automorphic Forms and Representations, Cambridge Univ. Press. 1997.

§4.3 德林费尔德证明函数域上二维朗兰兹局部猜想(1978年)

1974 年～1977 年,前苏联数学家德林费尔德(V. G. Drinfeld)类比于复椭圆曲线看成复平面对某个格的商空间,并且可由模函数参数化的理论,在函数域上研究一种新型的模结构. 他称之为"椭圆模",后人把这种模结构推广到高维,统称为德林费尔德模(椭圆模是其秩为 2 的情形). 利用这种椭圆模的精细结构和分类,德林费尔德得到$GL_2(A_K)$上足够多的自守表示,其中 K 是任意函数域. 这使他能够证明伽罗华群 $\mathrm{Gal}(K^{ac}/K)$的二维不可约表示均可来源于$GL_2(A_K)$的自守表示. 换句话说,他对函数域的情形证明了二维的朗兰兹猜想. 1988 年,他对于函数域上 $GL(2)$的自守形式证明了关于系数估计的 Petersson 猜想(我们在下节将知道,对于数域的情形,这个猜想是 1974 年由德林证明的,它是德林证明有限域上高维韦依猜想的推论). 德林费尔德的这些工作以及他在量子群和数学物理方面的工作获得 1990 年菲尔兹奖.

关于德林费尔德模的一般理论可参见 Deligne 和 Husemöller合写的介绍文章[4]. 在这里,我们仍想按历史的渊源和它与数论发展的联系加以叙述.

先让我们再回到希尔伯特第 12 问题:对于任意数域 K,如何明显地构作它的最大阿贝尔扩张 K^{ab}? 19 世纪末,Kronecker 和

Weber 解决了最基本情形，即对于有理数域 $\boldsymbol{Q}$，$\boldsymbol{Q}^{ab}$是所有分圆域的合成，即 $\boldsymbol{Q}^{ab}$是将所有复单位根（乘法群 $\boldsymbol{C}^{\times}$ 的所有有限阶元素）添加到 $\boldsymbol{Q}$ 中而得到的域. 我们知道，与有理数域相类比的函数域是有理函数域 $k=\boldsymbol{F}_q(T)$. 很自然会提出：k 的最大阿贝尔扩张 k^{ab} 如何明显地构作出来？这个问题一直到 1974 年才由 D. Hayes 给出解答. 他构作了一批"分圆"函数域，用函数域上的类域论证明了将这些分圆函数域合成就得到 k^{ab}. Hayes 的工作得益于他的老师 L. Carlitz 在 1935 年前后的研究. 在那个年代，函数域上数论研究的主流方向是 Artin 和 Hasse 等人关于 zeta 函数的黎曼猜想（即 1941 年的韦依猜想），Carlitz 的工作并未引起足够的注意.

有理函数域 $k=\boldsymbol{F}_q(T)$有一批显然的阿贝尔扩张，即常数域扩张 $k_n=\boldsymbol{F}_{q^n}(T)(n\geqslant 1)$. 伽罗华群 $\mathrm{Gal}(k_n/k)$同构于 $\mathrm{Gal}(\boldsymbol{F}_{q^n}/\boldsymbol{F}_q)$（$n$ 阶循环群）. Carlitz 用一种独特的方法得到 k 的一批新的阿贝尔扩域. 令 k^{ac}为 k 的代数闭包，$R=\boldsymbol{F}_q[T]$（这是整数环 $\boldsymbol{Z}$ 的模拟），Carlitz 给出 k^{ac}的一种新的 R-模结构. 以 $\mathrm{End}(k^{ac})$表示 k^{ac}的 $\boldsymbol{F}_q$-线性自同态构成的环. 则其中有如下的 Frobenius 自同构 $F\in\mathrm{End}(k^{ac})$，

$$F: k^{ac}\to k^{ac}, F(\alpha)=\alpha^q.$$

现在对于 $a\in\boldsymbol{F}_q$ 和 T 定义 ρ_a 和 $\rho_T\in\mathrm{End}(k^{ac})$，其中

$$\rho_a(\alpha)=a\alpha,$$

$$\rho_T(\alpha)=(T+F)(\alpha)=T\alpha+\alpha^q \qquad (\alpha\in k^{ac}).$$

一般地，对 $R=\boldsymbol{F}_q[T]$中每个多项式，

$$M(T)=a_dT^d+a_{d-1}T^{d-1}+\cdots+a_0 \qquad (a_i\in\boldsymbol{F}_q).$$

定义 $\rho_M\in\mathrm{End}(k^{ac})$，其中

$$\rho_M(\alpha)=M(T+F)(\alpha)=a_d\rho_{T^d}(\alpha)+a_{d-1}\rho_{T^{d-1}}(\alpha)+\cdots+a_1\rho_T(\alpha)+a_0\alpha,$$

其中 $\rho_{T^n}=\rho_{T^{n-1}}(\rho_T(\alpha))$.

例如

$$\rho_{T^2}(\alpha)=\rho_T(\rho_T(\alpha))=\rho_T(T\alpha+\alpha^q)$$

$$=T(T\alpha+\alpha^q)+(T\alpha+\alpha^q)^q$$
$$=T^2\alpha+(T+T^q)\alpha^q+\alpha^{q^2}.$$

由此给出 k^{ac}上一个 R-模结构(叫作 Carlitz 模)

$$\rho: R\to \mathrm{End}(k^{ac}), M\mapsto \rho_M.$$

对于 R 中每个首 1 多项式 $M=M(T)$,定义

$$\Lambda_M=\{\alpha\in k^{ac}\mid \rho_M(\alpha)=0\}.$$

这是 k^{ac}的一个 R-子模,类比于 m 次复单位根乘法群 $\mu_m=\{\alpha\in \boldsymbol{C}\mid \alpha^m=1\}$. 可以证明 Λ_M 是循环 R-模(类比于 μ_m 是循环群),并且有 R-模同构

$$\Lambda_M\cong R/(M)=\frac{\boldsymbol{F}_q[T]}{(M)}$$

(类比于 μ_m 同构于加法群 $\boldsymbol{Z}/(m)$). Λ_M 中元素叫作 M-torsion 元素(类比于 $\boldsymbol{C}$ 中 m 次单位根),循环 R-模 Λ_M 的每个生成元都叫作本原 M-torsion 元素(类比于 m 次本原单位根). 于是对每个本原 M-torsion 元素 λ,$\Lambda_M=\{\rho_A(\lambda)\mid A\in R/(M)\}$,并且 $\rho_A(\lambda)$本原当且仅当$(A,M)=1$(即 $A\in (R/(M))^\times$). 记 $\Phi(M)=|(R/(M))^\times|$,这是通常欧拉函数 $\varphi(m)=|(\boldsymbol{Z}/m\boldsymbol{Z})^\times|$的模拟. 类似于 $\varphi(m)$的公式,用中国剩余定理可知

$$\Phi(M)=|M|\cdot\prod_{P\mid M}\left(1-\frac{1}{|P|}\right),$$

其中 P 过 M 的所有首 1 不可约多项式因子,而$|M|=q^{\deg M}=|R/(M)|$.

将 Λ_M 中所有元素添加到 k 上而得的扩域表成 $k(\Lambda_M)$. 取定一个本原 M-torsion 元素 λ,则 Λ_M 中每个元素 $\rho_A(\lambda)$都是 λ 的多项式(系数属于 R),于是 $k(\Lambda_M)=k(\lambda)$. Carlitz 证明了:$k(\Lambda_M)/k$ 是伽罗华扩张,并且伽罗华群 $G=\mathrm{Gal}(k(\Lambda_M)/k)$同构于$(R/(M))^\times$,

$$(R/(M))^\times\simeq \mathrm{Gal}(k(\Lambda_M)/k),\qquad A\mapsto \sigma_A,$$

其中 σ_A 是由 $\sigma_A(\lambda)=\rho_A(\lambda)$所决定的自同构. 这类比于分圆数域 $\boldsymbol{Q}(\mu_m)=\boldsymbol{Q}(\zeta_m)$,$\mathrm{Gal}(\boldsymbol{Q}(\zeta_m)/\boldsymbol{Q})\cong(\boldsymbol{Z}/m\boldsymbol{Z})^\times$,于是 $k(\Lambda_M)$为 k 的阿

贝尔扩张. 进而，$R=\boldsymbol{F}_q[T]$中每个首 1 不可约多项式 $P=P(T)$（它生成 R 的非零素理想）在 $K=k(\Lambda_m)$中的分解情形也与素数 P 在 $\boldsymbol{Q}(\zeta_m)$中的分解规律非常相似. 基于这些事实，函数域 $k(\Lambda_m)$叫作**分圆函数域**. 将这些分圆域合在一起得到 $k_T=\bigcup_M k(\Lambda_M)$，这是 k 的（无限）阿贝尔扩张. 这个域是依赖于 $k=\boldsymbol{F}_q(T)$中超越元素 T 的选取. 若取另一个超越元素 $t=\dfrac{1}{T}$，按同样方式又可构作出 k 的另一个无限阿贝尔扩张 k_t. 再考虑常数域扩张 $\boldsymbol{F}_{q\infty}k=\boldsymbol{F}_{q\infty}(T)$，其中 $\boldsymbol{F}_{q\infty}$表示 $\boldsymbol{F}_q$ 的代数闭包，Hayes 于 1974 年证明了这三个扩张合在一起就是 k 的最大阿贝尔扩张，即

$$k^{ac}=\boldsymbol{F}_{q\infty}k_Tk_t.$$

Hayes 的这项工作几乎是与 Drinfeld 的椭圆模工作同时进行的. 当有了任意高维 Drinfeld 模的一般概念之后，Drinfeld 的椭圆模就是秩 1 的 Drinfeld 模，而 Hayes 使用的 Carlitz 模是秩 1Drinfeld 模的特殊情况. 1980 年，Hayes 考虑任意函数域 K 上的秩 1 Drinfeld 模，证明在添加此模的某些 torsion 点之后可得到 K 的最大阿贝尔扩域 K^{ab}. 换句话说，利用秩 1 Drinfeld 模，Hayes 对每个函数域都明显构作出它的最大阿贝尔扩域. 而对数域情形，至今只对有理数域和虚二次数域才解决了这个希尔伯特第 12 问题.

在这些工作的刺激之下，从 1980 年起对于分圆函数域和任意函数域上 Drinfeld 模的结构和分类研究形成一个高潮. 利用分圆函数域中的素理想分解规律和韦依 zeta 函数，得到了分圆函数域的类数解析公式，它很像分圆数域的类数公式. 作为 Bernoulli 数的模拟，得到两类“Bernoulli 元素”. 一类是 Carlitz 在 20 世纪 30 年代给出的，是 $\boldsymbol{F}_q(T)$中有理函数，用来判别“实”分圆函数域类数整除性. 另一类是 D. Goss 于 1982 年得到的，它们是 $R=\boldsymbol{F}_q[T]$中的一批多项式，用来判别分圆函数域除子类数整除性，这些结果是库默尔关于分圆数域类数整除性的结果到分圆函数域的成功移植. Goss 在构作这些 Bernoulli 多项式时是基于他引进的一种新

型 zeta 函数. 上述两类 Bernoulli 元素之间是否有联系至今仍是秘密，主要困难是 Goss zeta 函数没有适当的函数方程. 分圆数域的许多近代和现代结果（如分圆单位理论等）都已建立在函数域上，并且这些工作常常不是简单的类比和移植，在函数域中常常要采取与数域不同的方法，并且会出现与数域情形很不相同的现象. 某些问题在函数域情形有相当的困难. 比如说，目前在函数域情形还没有建立起满意的 Iwasawa 理论. 另一方面，由于任何函数域上都有 Drinfeld 模理论，使得在数域情形下只对分圆域建立的理论，在函数域情形可以走得更远. 例如，Hayes 使用“椭圆单位”对任意函数域都得到类数解析公式. 另一个精彩的例子是关于 zeta 函数特殊值的超越性问题. 对于黎曼 zeta 函数 $\zeta(s)=\sum_{n\geq 1} n^{-s}$，我们知道它在 $s=0$ 和负整数处值均是有理数（Bernoulli 数），而在正偶数处取值 $\zeta(2n)(n\geq 1)$ 是 π^{2n} 乘以有理数. 但是对正奇数 $2n+1(n\geq 1)$，$\zeta(2n+1)$ 的值至今是一个谜. 人们猜想它们都是超越数，直到 1978 年 Apéry 才证明了 $\zeta(3)$ 是无理数. 但是对函数域情形的 Goss zeta 函数，利用 Drinfeld 模理论，于靖关于 Goss zeta 函数 $\zeta_\infty(s)$ 在整数处取值给出了圆满的超越性结果. 首先指出，在函数域中 $q-1$ 是数域情形数 2 的类比，因为 2 是整数环 $\boldsymbol{Z}$ 中可逆元 ± 1 的个数，而 $q-1$ 是 $R=\boldsymbol{F}_q[T]$ 中可逆元 $\boldsymbol{F}_q^\times$ 的个数. 其次，Carlitz 给出指数函数 e^{2iz} 在函数域中的模拟 $e(z)$，$e(z)$ 是周期函数，周期 $\bar{\pi}$ 为 $k=\boldsymbol{F}_q(T)$ 上的超越元素（类比于 e^{2iz} 的周期 π 是 $\boldsymbol{Q}$ 上的超越数）. 对于整数 n，当 $q-1\mid n$ 时，$\zeta_\infty(n)/\bar{\pi}^n$ 属于 k，而于靖证明了：当 $q-1\nmid n$ 时，$\zeta_\infty(n)$ 均是 k 上的超越元素.

函数域 K 有两种类数，一个是零次除子类数 $h(K)$，另一个是理想类数 $h(O_k)$，这里 O_K 表示 K 的整元素环，即 K 中对 $R=\boldsymbol{F}_q[T]$ 整的元素构成的狄德金整环. $h(K)=R(K)h(O_K)$，其中 $R(K)$ 是域 K 的 regulator，由单位群 $U_K=O_K^\times$ 的一组基所决定. 数域情形 $R(K)$ 是很难计算的实数，函数域情形 $R(K)$ 也很难计算，

但它是正整数. 现在我们对分圆函数域 $K=k(\Lambda_M)$ 推导除子类数 $h(K)$ 的计算公式. 根据韦依的理论, K 的韦依 zeta 函数为

$$Z_K(U)=\sum_A U^{\deg A}=\frac{F(U)}{(1-U)(1-qU)} \quad (U=q^{-s}),$$

其中 A 过 K 的所有整除子, $F(U)\in \mathbf{Z}[U]$, 而 $h(K)=F(1)$. $k=\boldsymbol{F}_q(T)$ 中的除子 $\infty=\left(\dfrac{1}{T}\right)$ 到 K 的扩充都叫作 K 的无限除子. 以 λ 表示一个本原 M-torsion 元素, 则 $K=k(\lambda)$, 利用简单的 ∞-adic 分析可以证明 ∞ 在 K 中的分解域是 $K^+=k(\lambda^{q-1})$, 即 ∞ 在 K^+ 中为 g 个不同素除子乘积 $\infty=\mathscr{P}_1\cdots\mathscr{P}_g$, 其中 $g=[K^+:k]=[K:k]/[K:K^+]=\Phi(M)/(q-1)$. 而每个 $\mathscr{P}_i$ 在 K 中均完全分歧: $\mathscr{P}_i=P_i^{q-1}(1\leqslant i\leqslant g)$. 于是 $\deg P_i=1(1\leqslant i\leqslant g)$. 我们把 K^+ 叫作 K 的最大"实"子域. 如果把 $\mathrm{Gal}(K/k)$ 等同于 $(R/(M))^\times$, 则 $\mathrm{Gal}(K/K^+)$ 等同于 $(R/(M))^\times$ 的子群 $\boldsymbol{F}_q^\times$, 从而 $\mathrm{Gal}(K^+/k)$ 等同于 $(R/(M))^\times$ 对于 $\boldsymbol{F}_q^\times$ 的商群. 群 $(R/(M))^\times$ 的复特征 $\chi:(R/(M))^\times\to\boldsymbol{C}^\times$ 叫作多项式环 $R=\boldsymbol{F}_q[T]$ 模 M 的狄里赫利特征. 这 $\Phi(M)$ 个特征叫作 K 的特征. 而域 K^+ 的特征则指的是 $\mathrm{Gal}(K^+/k)$ 的特征, 由上所述它们相当于满足条件 $\chi(\boldsymbol{F}_q^\times)=1$ 的那些模 M 狄里赫利特征, 这种特征叫**实特征**, 共有 $\dfrac{\Phi(M)}{q-1}=[K^+:k]$ 个. 于是我们有

$$\begin{aligned}Z_K(U)&=\sum_A U^{\deg A}=\prod_P(1-U^{\deg P})^{-1}\\&=\prod_{P\mid\infty}(1-U^{\deg P})^{-1}\cdot\prod_{P\nmid\infty}(1-U^{\deg P})^{-1},\end{aligned}$$

其中 P 过 K 的所有素除子. 由上述 ∞ 在 K 中分解特性可算出上式右边第一个因子

$$\prod_{P\mid\infty}(1-U^{\deg P})^{-1}=(1-U)^{-g},$$

而第二部分因子中 P 过 O_K 的所有非零素理想, 即

$$\prod_{P\nmid\infty}(1-U^{\deg P})^{-1}=\sum_{A\subseteq O_K}U^{\deg A}=\sum_{A\subseteq O_K}NA^{-S}$$
$$(NA=q^{\deg A}).$$

这正是域 K 的狄德金 zeta 函数. 为了计算它,我们需要 k 的每个有限素除子(即 R 中首 1 不可约多项式)$p=p(T)$在分圆函数域 $K=k(\Lambda_M)$中的分解规律. 可以和数域情形相仿地证明:若 $M=p^lN$,其中$(N,p)=1,l\geqslant 0$. 则

$$pO_K=(P_1\cdots P_g)^e,\deg P_i=f\quad(1\leqslant i\leqslant g),$$

其中 $e=\Phi(p^l)$,f 为 p 模 N 的乘法阶,$g=\Phi(N)/f$. 然后便可跟数域情形一样证得

$$\sum_{A\subseteq O_K}U^{\deg A}=\prod_{P\nmid\infty}(1-U^{\deg P})^{-1}=\prod_{\chi}L(U,\chi^*),$$

其中 χ 过模 M 的所有狄里赫利特征,χ^* 是结合于 χ 的本原特征,而 L-函数 $L(U,\chi^*)$定义为

$$L(U,\chi^*)=\prod_{p}(1-\chi^*(p)U^{\deg p})^{-1}=\sum_{A\in R_1}\chi^*(A)U^{\deg A}$$
$$=\sum_{n=0}^{\infty}\sigma_n(\chi^*)U^n.\tag{1}$$

这里 p 过 $R=\boldsymbol{F}_q[T]$中所有首 1 不可约多项式,A 过 R 中所有首 1 多项式(它们形成的集合记为 R_1),而

$$\sigma_n(\chi^*)=\sum_{\substack{A\in R_1\\ \deg A=n}}\chi^*(A).$$

综合上述我们得到

$$\frac{F(U)}{(1-U)(1-qU)}=Z_k(U)=(1-U)^{-g}\prod_{\chi}L(U,\chi^*),$$
$$g=\Phi(M)/(q-1).$$

当 $\chi=1$(平凡特征)时,$L(U,\chi^*)$就是 $k=\boldsymbol{F}_q(T)$的狄德金 zeta 函数

$$\sum_{A\in R_1}U^{\deg A}=\sum_{n=0}^{\infty}q^nU^n=(1-qU)^{-1}.$$

于是

$$F(U)=\prod_{\chi\neq 1}L(U,\chi^*)/(1-U)^{g-1}. \tag{2}$$

当 $\chi\neq 1$ 时，以 D_χ 表示 χ^* 的导子，$D_\chi|M$，令 $d_\chi=\deg D_\chi$. 可以证明当 $n\geqslant d_\chi$ 时，$\sigma_n(\chi^*)=0$. 从而由(1)式可知 $L(U,\chi^*)$是关于 U 的多项式. 又若 χ 是实的非平凡特征时，$L(1,\chi^*)=0$，从而多项式 $L(U,\chi^*)$有因子$(1-U)$. 于是(2)式可写成(注意非平凡实特征共 $g-1$ 个)

$$F(U)=\prod_{\chi\text{非实}}L(U,\chi^*)\cdot\prod_{\substack{\chi\text{实}\\ \chi\neq 1}}\left(\frac{L(U,\chi^*)}{1-U}\right).$$

当 χ 是模 M 非实特征时，

$$L(1,\chi^*)=\sum_{n=0}^{d_\chi-1}\sigma_n(\chi^*).$$

而当 χ 是非平凡实特征时，

$$\begin{aligned}L(1,\chi^*)&=\lim_{U\to 1}\frac{L(U,\chi^*)}{1-U}\\&=\lim_{U\to 1}\left(\sum_{n=0}^{d_\chi-1}\sigma_n(\chi^*)U^n\right)/(1-U)\\&=-\sum_{n=1}^{d_\chi-1}n\sigma_n(\chi^*).\end{aligned}$$

于是最后得到分圆函数域 $K=k(\Lambda_M)$除子类数 $h(K)$的解析公式

$$h(K)=F(1)=h^+\cdot h^-,$$

其中

$$h^+=\prod_{\substack{\chi\text{实}\\ \chi\neq 1}}h(\chi),\qquad h(\chi)=-\sum_{n=1}^{d_\chi-1}n\sigma_n(\chi^*),$$

$$h^-=\prod_{\chi\text{非实}}h(\chi),\qquad h(\chi)=\sum_{n=0}^{d_\chi-1}\sigma_n(\chi^*).$$

类似可证 h^+ 即是最大实子域 K^+ 的除子类数，于是 h^+ 为正整数而 $h^-=h/h^+$ 叫作 K 的相对类数. 由于 $h(\chi)$都是代数整数而 $h^-=$

$h/h^+\in \boldsymbol{Q}$,可知 h^- 也是正整数(数域情形证明此事要困难一些).

K 和 K^+ 整数环的单位群都是秩 $g-1$ 的有限生成交换群. 和数域一样,寻找一组基(基本单位系)是困难的. 但是可以找到一批单位 $\frac{\rho_A(\lambda)}{\lambda}(A\in(R/(M))^\times, A\neq 1)$,这些叫**分圆单位**. 当 M 是不可约多项式的方幂时,它们生成的分圆单位群 C_K 对 O_{K^+} 的单位群 U_K^+ 指数有限,并可算出 $[U_K^+:C_K]=h_K^+$. 这些结果都与数域情形相仿.

1983 年 Goss 引进一种新型的 zeta 函数

$$\zeta_\infty(s)=\sum_{A\in R_1}A^{-s}=\sum_{n=0}^{\infty}a_n(s),\quad a_n(s)=\sum_{\substack{A\in R\\ \deg A=n}}A^{-s}.$$

自变量 s 取值为整数,而函数值属于 $k=\boldsymbol{F}_q(T)$ 对于无限素除子 $\infty=\left(\frac{1}{T}\right)$ 的完备化 $k_\infty=\boldsymbol{F}_q=\left(\left(\frac{1}{T}\right)\right)$(它类比于 $\boldsymbol{R}=\boldsymbol{Q}_\infty$). 当 s 为正整数对定义 $\zeta_\infty(s)$ 的级数对于 k_∞ 中的∞-adic 拓朴收敛,从而取值于 k_∞. 对于负整数 $-m$, $\zeta_\infty(-m)$ 是多项式

$$\zeta_\infty(-m)=\sum_{\substack{A\in R_1\\ \deg A\leqslant\frac{m}{q-1}}}A^m\in\boldsymbol{F}_q[T].$$

对于正整数 i,当 $(q-1)\nmid i$ 时,令

$$\beta_i(T)=\zeta_\infty(-i)=\sum_{\substack{A\in R_1\\ \deg A\leqslant\frac{i}{q-1}}}A^i.$$

而当 $(q-1)\mid i$ 时,$\zeta_\infty(-i)=0$,我们定义其"微商"为

$$\beta_i(T)=-\sum_{\substack{A\in R_1\\ \deg A\leqslant\frac{i}{q-1}}}A^i\deg A.$$

这些 $\beta_i(T)$ 均是 $\boldsymbol{F}_q[T]$ 中多项式. 它们就是 Goss 引进的"贝努利多项式",因为不但有类似于贝努利数的递推关系和同余性质,而且有库默尔关于分圆数域类数整除性结果的如下模拟:设 p 是 $\boldsymbol{F}_q$ 的特征,$K=k(\Lambda_P)$,其中 P 是 $\boldsymbol{F}_q[T]$ 的 d 次首 1 不可约多项式,

h^+和h^-分别为K^+的除子类数和K的相对类数，则$p|h^+$当且仅当存在正整数$i\leqslant q^d-2,(q-1)|i$，使得$P|\beta_i(T)$. 而$p|h^-$当且仅当存在正整数$i\leqslant q^d-2,(q-1)\nmid i$使得$P|\beta_i(T)$.

现在介绍Carlitz在1935年～1941年定义的域$k=\boldsymbol{F}_q(T)$上另一批"贝努利元素". 通常贝努利数B_n是由

$$\frac{t}{e^t-1}=\sum_{n\geqslant 0}\frac{B_n}{n!}t^n$$

定义的. 所以Carlitz首先探讨指数函数的阶乘$n!$应当如何推广到函数域上. 指数函数e^t是以$2\pi i\boldsymbol{Z}$为周期的复值函数，$2\pi i\boldsymbol{Z}$是$\boldsymbol{C}$中秩1的格. 我们知道，$\boldsymbol{Z},\boldsymbol{Q},\boldsymbol{R}$和$\boldsymbol{C}$在函数域中的模拟分别为$R=\boldsymbol{F}_q[T],k=\boldsymbol{F}_q(T),k_\infty=\boldsymbol{F}_q((\frac{1}{T}))$和$k_\infty^{ac}$($k_\infty$的代数闭包). 对每个非零元素$\gamma\in k_\infty^{ac}$，$\gamma R$是$k_\infty^{ac}$中的秩1格(离散加法子群)，可以如下构作一个以$L=\gamma R$为周期格的函数

$$e_L(z)=z\prod_{0\neq a\in L}\left(1-\frac{z}{a}\right)=\lim_{h\to\infty}z\prod_{0\neq a\in L(h)}\left(1-\frac{z}{a}\right),$$

其中$L(h)=\{\gamma A\mid A\in R,\deg A<h\}$. 可以证明当$z\in k_\infty^{ac}$时，右边极限在$k_\infty^{ac}$中收敛，从而$e_L(z)$是从$k_\infty^{ac}$到$k_\infty^{ac}$的函数，并且易知$e_{\gamma R}(\gamma z)=\gamma e_R(z)$. Carlitz发现有一个特定的秩1格$L$，使得$e_L(z)$对于展开有简单表达式. 他引进如下的符号

$$[i]=T^{q^i}-T\ (i\geqslant 1),$$
$$D_0=1,$$
$$D_i=[i]D_{i-1}^q=[i][i-1]^q[i-2]^{q^2}\cdots[1]^{q^{i-1}}\quad(i\geqslant 1).$$

通过∞-adic分析可知$\prod_{i=1}^{\infty}\left(1-\frac{[i]}{[i+1]}\right)$收敛为$k_\infty$中元素，表示成$\mu$. 以$\upsilon$表示方程$y^{q-1}=-[1]$在$k_\infty^{ac}$中的一个解，令$\tilde{\pi}=\mu\upsilon\in k_\infty^{ac}$. 则对于秩1格$L=\tilde{\pi}R$，Carlitz计算出

$$e_L(z)=e_{\tilde{\pi}R}(z)=\sum_{i=0}^{\infty}\frac{z^{q^i}}{D_i},\tag{1}$$

并且有函数方程

$$e(Tz)=Te(z)+e(z)^q=\rho_T(e(z)).$$

又由(1)式易知 $e_L(z+z')=e_L(z)+e_L(z')$ 和 $e_L(az)=ae_L(z)(a\in \boldsymbol{F}_q)$. 由这些关系可知对每个 $M\in R$,均有 $e_L(Mz)=\rho_M(e_L(z))$. 其中 ρ_M 是 Carlitz 作用. 换句话说,我们有交换图表

$$\begin{array}{ccc} k_\infty^{ac} & \xrightarrow{e_L} & k_\infty^{ac} \\ M\downarrow & & \downarrow \rho_M \\ k_\infty^{ac} & \xrightarrow{e_L} & k_\infty^{ac} \end{array} \qquad (\text{对每个 } M\in R).$$

左边是 R 在 k_∞^{ac} 上的自然作用 $z\mapsto Mz$,即 k_∞^{ac} 看成为自然 R-模. 通过指数映射 e_L 之后,k_∞^{ac} 变成 Carliltz R-模:$z\mapsto \rho_M(z)$. 这就是 Carlitz 模的来源! 30 年之后,Drinfeld 考虑 k_∞^{ac} 的秩 2 格 $L=\alpha R\oplus\beta R(\alpha,\beta\in k_\infty^{ac})$. 类似定义以 L 为周期格的指数函数 $e_L(z)$,并且有映射 $D:R\to\mathrm{End}(k_\infty^{ac}),M\mapsto D_M$,使得有如下的交换图表

$$\begin{array}{ccc} k_\infty^{ac} & \xrightarrow{e_L} & k_\infty^{ac} \\ M\downarrow & & \downarrow D_M \\ k_\infty^{ac} & \xrightarrow{e_L} & k_\infty^{ac} \end{array}$$

即对每个 $M\in R,e_L(Mz)=D_M(e_L(z))$. 从而对于 k_∞^{ac} 的每个秩 2 格 L,k^{ac} 都得到一个 R-模结构 D,M 在 $z\in k_\infty^{ac}$ 上的作用为 $D_M(z)$. 由于 e_L 以 L 为周期格,所以实际上 e_L 是从 k_∞^{ac}/L 到 k_∞^{ac} 的 $\boldsymbol{F}_q$-线性映射(并且是满射). 回忆数域情形,$\boldsymbol{C}$ 对于一个秩 2 格 L 的商 $\boldsymbol{C}/L$ 就是一条复椭圆曲线,所以 Drinfeld 把这种模结构叫作椭圆模. 对于秩 1 情形,k_∞^{ac} 的任意两个秩 1 模 $L_1=\gamma_1R$ 和 $L_2=\gamma_2R$ 都是相似的:$L_1=\frac{\gamma_1}{\gamma_2}L_2$. 可知它们本质上给出 k_∞^{ac} 的同一个模结构,即 k_∞^{ac} 上所有秩 1 Drinfeld 模本质上只有一个,就是 Carlitz 模. 对于秩 2 情形,k_∞^{ac} 的秩 2 格有无穷多相似等价类,从而给出 k_∞^{ac} 的无穷多本质不同的椭圆模结构. Drinfeld 就是用这样丰富的椭圆模结构给出充分多的自守表示,从而证明了 $n=2$ 情形在函数域上的朗兰兹猜想. 再进一步,对于数域情形 $\boldsymbol{C}$ 是 $\boldsymbol{R}$ 的 2 次扩张,从而 $\boldsymbol{C}$ 中格的秩

只能为 1 和 2. 但是 k_∞^{ac} 为 k_∞ 的无限次扩张，所以对每个正整数 n，k_∞^{ac} 中均有秩 n 的格，从而 k_∞^{ac} 上均有（许多）秩 n 的 Drinfeld 模结构. 用它们应当构作出函数域上 $GL(n)$ 的许多自守表示. 但目前对 $n\geqslant 3$ 的情形，秩 n Drinfeld 模的分类问题（moduli，参量空间）没有完全被解决，所以函数域上 $n\geqslant 3$ 的朗兰兹猜想还未被证明.

现在回头来介绍 Carlitz 给出的阶乘 $n!$ 在函数域中的模拟. 设正整数 n 有 q-adic 展开

$$n=a_0+a_1q+\cdots+a_sq^s \quad (0\leqslant a_i\leqslant q-1).$$

定义 $\boldsymbol{F}_q[T]$ 中的多项式

$$\Gamma_n=D_0^{a_0}D_1^{a_1}\cdots D_s^{a_s},$$

其中符号 D_i 的定义如前所述. 我们知道 $n!$ 有素因子分解式 $n=\prod_p p^{e_p}$，其中 $e_p=\sum_{m\geqslant 1}\left[\frac{n}{p^m}\right]$. 类似地可证 Γ_n 在 $\boldsymbol{F}_q[T]$ 中有素因子分解式

$$\Gamma_n=\prod_P P^{e_P}, e_p=\sum_{m\geqslant 1}\left[\frac{n}{q^{m\deg P}}\right].$$

由此和其他一些理由可把 Γ_n 看作是 $n!$ 的模拟. 然后类比于 $\frac{t}{e^t-1}=\sum_{n\geqslant 1}\frac{B_n}{n!}t^n$，Carlitz 由下式定义了一批 $\boldsymbol{F}_q(T)$ 中元素 B_n，

$$\frac{z}{e_{\tilde{\pi}R}(z)}=\sum_{n\geqslant 0}\frac{B_n}{\Gamma_n}z^n.$$

这些有理函数 $B_n\in\boldsymbol{F}_q(T)$ 的部分分式展开和贝努利数的部分分式展开相类似，并且若 $(q-1)\nmid n$ 时 $B_n=0$，而当 $(q-1)|n$ 时，$B_n=\Gamma_n\zeta_\infty(n)/\tilde{\pi}^n$. 1991 年 Okada 给出 B_n 和 $K^+=k(\Lambda_P)^+$ 的理想类数 $h=h(O_K^+)$ 之间的如下联系：若 P 是 $\boldsymbol{F}_q[T]$ 中 d 次不可约首 1 多项式，p 是 q 的素因子. 如果 $p|h$，则存在正整数 $n\leqslant q^d-2$，$(q-1)|n$，使得 P 除尽 B_n 的分子多项式. 这个结果使得 B_n 更有资格作为贝努利数的模拟.

以上我们介绍了分圆函数域和 Drinfeld 模理论的初步内容. 我们又一次看到数域和函数域研究的互动和促进. 代数数论起步时,许多方面受函数论的启发. 例如素数在数域中的分歧概念来源于代数函数在某点处的分歧性质. Hensel 的 p-adic 数和局部域来源于函数在局部的 Taylor 展开. 在赋值理论建立起来之后,从 20 世纪 20 年代起,函数域研究开始借用数域成果,发展了以韦依 zeta 函数和韦依定理为代表的算术理论.

到了 20 世纪 50 年代,类比于函数域的有限常数域扩张,Iwasawa 建立起数域的近代分圆域理论. 在 Drinfeld 模理论指导下,数域的分圆理论又转移到函数域中,促进了函数域算术研究的新进展.

关于分圆函数域和 Drinfeld 模的进一步理论可见参考文献(以及这些文献所引用的文章). Drinfeld 模的研究被用于研究函数域上的模形式,自守形式和自守表示,也用来研究特征 p 的代数簇(阿贝尔簇和 Shimura 簇)和函数域上高维朗兰兹猜想.

参 考 文 献

[1] G. Anderson, Rank one elliptic A-modules and A-harmonic series, Duke Math. J., 73(1994), 491～542.

[2] L. Carlitz, On certain functions connected with polynomials in a Galois field, Duke Math. J., 1(1935), 137～168.

[3] —, An analogue of Bernoulli polynomials, Duke Math. J., 8 (1941), 405～412.

[4] P. Deligne and D. Husemöller, Survey of Drinfeld modules, Contemp. Math. 67(1987), 25～91.

[5] V. G. Drinfeld, Elliptic modules, Math. Sbornik 94(1974), 594~627 II, 31(1977), 159~170.

[6] —, Langlands conjecture for GL(2) over function fields, Proc. Internat. Congress Math., Helsinki, 1978.

[7] —, Two-dimensional l-adic represen tations of the fundamental group of a curve over a finite field and automorphic forms on GL(2), Amer. J. Math. 105(1983), 85~114.

[8] —, The proof of Petersson's conjecture for GL(2) over a global field of characteristic p, Funct. Anal. and Appli. 22 (1988), 28~43.

[9] S. Galovich and M. Rosen, Units and class groups of cyclotomic function fields, J. Number Theory 14(1982), 156~184.

[10] E.-U. Gekeler, Drinfeld Modular Curves, Lecture Notes in Math. 1231, Springer-Verlag, 1986.

[11] D. Goss, Kummer and Herbrand criterion in the arithmetic of function fields, Duke Math. J. 49(1982), 377~384.

[12] —, On a new type of L-function associated to algebraic curves over finite fields, Pacific J. Math. 105(1983), 143~181.

[13] —, Basic Structures of Function Field Arithmetic, Springer-Verlag, 1996.

[14] G. Harder and D. Kazhdan, Automorphic forms on GL(2) over function fields (after V. G. Drinfeld), in Automorphic Forms, Representations and L-functions, eds. A. Borel and W. Casselman. Amer. Math. Soc. (1979), 357~379.

[15] D. Hayes, Explicit class field theory for rational function field, Trans. Amer. Math. Soc. 189(1974), 77~91.

[16] —, Explicit class field theory in global function fields, in

Studies in Algebra and Number Theory 16(1980),173～217.

[17] —,A brief introduction to Drinfeld modules,in The Arithmetic of Function Fields, eds. D. Goss et al, de Gruyter (1992),1～32.

[18] B. Poonen, Local height functions and the Mordell-Weil theorem for Drinfeld modules,Compositio Math. 97(1995), 349～368.

[19] D. Thakur,On characteristic p zeta functions,Compositio Math. 99(1995),231～247.

[20] Jing Yu, Transcendence and Drinfeld modules, Invent. Math. 83(1986),507～517.

[21] —,Transcendence and Drinfeld modules: several variables, Duke Math. J. 58(1989),559～575.

§4.4 德林证明高维韦依猜想(1973年)

设C是有限域$\mathbf{F}_q$上一条不可约非奇异代数曲线. 以N_n表示C在$\mathbf{F}_{q^n}$中的射影点数,曲线C的zeta函数定义为

$$Z_C(U)=\exp\left(\sum_{n\geq 1}\frac{N_n}{n}U^n\right).$$

韦依在计算了大量例子之后,提出如下猜想:

(1) $$Z_C(U)=\frac{L(U)}{(1-U)(1-qU)},$$

$$L(U)=\sum_{i=1}^{2g}(1-\alpha_iU)\in\mathbf{Z}[U],$$

其中 g 为曲线 C 的亏格.

(2)（函数方程）$Z_C(U)=(\sqrt{q}U)^{2g-2}Z_C\left(\frac{1}{qU}\right)$.

(3)（黎曼猜想的模拟）$|\alpha_i|=\sqrt{q}\ (1\leqslant i\leqslant 2g)$.

(1)和(2)可以用黎曼-洛赫定理比较容易地证明(韦依用 Adele 语言给出黎曼-洛赫定理的一个纯代数证明,它适用于任何整体域上的代数曲线). 而对于猜想(3),E. Artin 首先对超椭圆曲线 $y^2=f(x)$(相当于二次函数域)的情形于 1924 年给出一些例子. Hasse 于 1933 年和 1934 年证明了椭圆曲线($g=1$)的情形. Hasse 给了两个不同的证明,第一个证明是考虑定义在 $\boldsymbol{F}_q$ 上椭圆曲线 E 的所有点 $E(\boldsymbol{F}_{q^\infty})$,其中 $\boldsymbol{F}_{q^\infty}$ 是 $\boldsymbol{F}_q$ 的代数闭包. 若(x,y)是曲线 E 的一个 $\boldsymbol{F}_{q^\infty}$-点,则$(x^q,y^q)$也是 E 的 $\boldsymbol{F}_{q^\infty}$-点. 于是映射$(x,y)\mapsto(x^q,y^q)$是阿贝尔群 $E(\boldsymbol{F}_{q^\infty})$的自同态,叫作 Frobenius 自同态 F. 而 F 的不动点即相当于 $\boldsymbol{F}_q$-点. 所以 E 上 $\boldsymbol{F}_q$-点个数即是映射 F 的不动点数. 利用复乘理论把映射 F 提升到特征零情形,如此证明了猜想(3)(对 $g=1$ 情形). 第二个证明则是研究椭圆曲线 E 的自同态环.

1940 和 1941 年,韦依在两篇短文中叙述了对任意亏格情形猜想(3)的两个证明思想. 第一个证明(1940 年)是把曲线到自身的 Frobenius 对应诱导出曲线的雅柯比簇上的一个自同态,从而这个对应结合一个 l-adic 的 $2g$ 阶方阵,而 Frobenius 对应的固定点数即是这个方阵的迹函数,然后由"Rosatti 对合"的正性推出猜想(3)(曲线的雅柯比簇是 g 维阿贝尔簇,g 为曲线的亏格. 当 $g=1$ 时,椭圆曲线 E 的雅柯比簇即是 E 自身,这就归结于 Hasse 情形). 第二个证明(1941 年)则完全依赖于 Severi 的曲线到自身对应理论,用对应的相交数来表达曲线的 zeta 函数. 然后韦依用这些相交数来定义所有对应组成的环上一个"迹函数",由这个迹函数的正性(即 Castelnuovo 不等式)证明猜想(3). 这两种证明思想本质上是 Hasse 对 $g=1$ 给出两个证明到一般情形的推广.

韦依这两个证明思想所用的基本工具(曲线对应理论和正性论断)是意大利代数几何学派已经准备好的现成工具，但是缺乏严格的数学语言. 这促使韦依写了《代数几何基础》一书(1946 年出版)，建立了抽象的代数几何理论，然后在 1948 年又写了两本书《阿贝尔簇和代数曲线》和《代数曲线和它诱导的代数簇》，分别严格地给出上述两个证明，并且把第一个证明推广到阿贝尔簇上.

1949 年，韦依对有限域上高维代数簇计算了一些例子(对角方程定义的超曲面，阿贝尔簇和 Grasmann 簇)，提出了如下的高维猜想：设 X 是 $\boldsymbol{F}_q$ 上非奇异的 n 维射影代数簇，N_m 为 X 在 $\boldsymbol{F}_{q^m}$ 上的射影点数，则

(A) X 的 zeta 函数

$$Z_X(U) = \exp\left(\sum_{m \geqslant 1} \frac{N_m}{m} U^m \right)$$

是 U 的有理函数.

(B) $$Z_X(U) = \frac{P_1(U)P_3(U)\cdots P_{2n-1}(U)}{P_0(U)P_2(U)\cdots P_{2n}(U)}, \tag{3.1}$$

其中

$$P_i(U) = \prod_{j=1}^{b_i}(1 - \alpha_{ij}U) \in \mathbf{Z}[U], b_i = \deg P_i(U),$$
$$|\alpha_{ij}| = q^{i/2} \quad (1 \leqslant j \leqslant b_i) \quad (0 \leqslant i \leqslant 2n)$$

(黎曼猜想的模拟).

(C) 将诸根 $\{\alpha_{ij}\}$ 的下标 j 作适当调整，可使 $\alpha_{ij}\alpha_{2n-i,j} = q^n$(对任意 i,j). 这相当于 $Z_X(U)$ 的一个函数方程.

(D) 设 X 可提升为复数域上非奇异射影代数簇 $\tilde{X}$，则 b_i 为 $\tilde{X}$(看成复流形时)经典上同调群 $H^i(\tilde{X},\boldsymbol{C})$ 的维数(第 i 个 Betti 数).

这个猜想的背景是紧复流形 $\tilde{X}$ 的代数拓朴学. 对于 $\tilde{X}$ 的任意自同态 $\mathscr{F}$，它可诱导出上同调群 $H^i(\tilde{X},\boldsymbol{C})$(这是有限维复向量空间)上的线性映射(仍记为 $\mathscr{F}$)以 #Fix($\mathscr{F}^n$)表示 $\mathscr{F}^n$ 在 $\tilde{X}$ 中的固

定点数，代数拓朴中有 Lefschetz 公式(其中 n 为 $\widetilde{X}$ 的复维数).

$$\#\mathrm{Fix}(\mathscr{F}^n) = \sum_{i=0}^{2n}(-1)^i T_r(\mathscr{F}^n | H^i(\widetilde{X}, \boldsymbol{C})). \qquad (3.2)$$

于是在形式上有

$$\exp\left(\sum_{m\geqslant 1}\frac{U^m}{m}\#\mathrm{Fix}(F^n)\right)$$
$$= \sum_{i=0}^{2n}\det(1 - U \cdot F | H^i(\widetilde{X}, \boldsymbol{C}))^{(-1)^{i+1}}. \qquad (3.3)$$

对于有限域 $\boldsymbol{F}_q$ 上非奇异射影代数簇 X，有 $X(\boldsymbol{F}_{q^\infty})$ 到自身的 Frobenius 映射 $F:(x_1,\cdots,x_n)\mapsto(x_1^q,\cdots,x_n^q)$. 而 F^m 的固定点个数 $\#\mathrm{Fix}(F^m)$ 就是 X 上的 $\boldsymbol{F}_{q^m}$-点个数 N_m. 因此

$$Z_X(U) = exp\left(\sum_{m\geqslant 1}\frac{U^m}{m}\#\mathrm{Fix}(F^n)\right).$$

韦依指出，要证明他的高维猜想，需要寻求对于有限域上代数簇合适的上同调理论，使得公式(3.2)成立. 于是 $Z_X(U)$ 可写成公式(3.3)右边的形式，由此就可推出猜想(B)中的有理函数表达式(3.1)和猜想(D)中对于 $\deg P_i$ 的论断，并且 $Z_X(U)$ 的函数方程应当来源于上同调群的对偶理论.

20 世纪 50 年代以来，复代数几何取得很大进步，发展了拓朴工具和层论工具，韦依本人从 1949 年起在芝加哥大学研究具有 Zariski 拓朴的抽象代数几何的纤维丛理论，1954 年前后 Serre 对于特征 p 的情形作了一系列重要的工作. 但是这些特征 p 的代数簇上同调理论不适合于证明韦依猜想，因为这些上同调空间都是特征 p 的向量空间，线性算子的迹是 $\boldsymbol{F}_p$ 中元素，所以只能给出有理点数的模 p 信息. 到了 1960 年，Dwork 证明了 $Z_X(U)$ 是 U 的有理函数. Dwork 采用了一套新的方法，使用 $\boldsymbol{Q}_p$ 上希尔伯特空间的泛函分析理论. Dwork 理论在 20 世纪后期得到很大发展，但是对于韦依猜想并没有取得进一步的突破. 尽管如此，它促使人们去寻求局部域 $\boldsymbol{Q}_l$ 上的上同调理论，因为这是特征零的上同调理论，并

且希望用所有局部信息来把握整体信息. 1964 年, M. Artin(E. Artin 的儿子)和 A. Grothendieck 对每个素数 l($l \neq \boldsymbol{F}_q$ 的特征 p), 都给出好的 l-adic 上同调理论. 根据这种理论, 对每个 $l \neq p$, 都可证明出

$$Z_X(U) = \prod_{i=0}^{2n} P_{i,l}(U)^{(-1)^{i+1}} = \frac{P_{1,l}(U)\cdots P_{2n-1,l}(U)}{P_{0,l}(U)\cdots P_{2n,l}(U)},$$

其中 $P_{i,l} \in \boldsymbol{Q}_l[U]$, 并且这些多项式 $P_{i,l}(U)$(在 $\boldsymbol{Q}_l$ 的代数闭包中)的根满足猜想(C). 而且当 X 提升为复流形 $\tilde{X}$ 时, $P_{i,l}(U)$ 的次数是 $\tilde{X}$ 的第 i 个 Betti 数 b_i. 但是他们不能证明所有多项式 $P_{i,l}(U)$ ($l \neq p$)均是同一个多项式 $P_i(U)$, 并且 $P_i(U)$ 的系数属于 $\boldsymbol{Q}$. 当然, 他们也不能证明黎曼猜想的模拟.

1973 年, 德林(P. Deligne)最终证明了整个韦依猜想, 他的这项工作获得 1978 年 Fields 奖, 他的工作主要体现在两个方面的突破性贡献.

1. 使用了 Lefschetz 束(pencils)的 Monodromy 群理论.

Lefschetz 于 1924 年对复代数簇的上同调理论做出一系列重要工作. 他把一个射影簇用超平面去截, 得到许多纤维, 而这个射影簇的上同调群可以用这些纤维的上同调群来表达. 这个理论转到 l-adic 上同调上去并不困难, 但是德林意识到: 要用于韦依猜想还需要证明每个维数的纤维组成的 Lefschetz 束都有很大的 Monodromy 群. 德林对于 l-adic 情形证明了这个结果, 由此得到诸多项式 $P_{il}(U)$ 是与 l 无关的多项式 $P_i(U)$, 并且 $P_i(U) \in \boldsymbol{Q}[U]$. 所以剩下的事情是证明黎曼猜想的模拟.

2. $Z_X(U)$ 与模形式的联系.

在韦依高维猜想提出后不久, 人们就意识到它和 Ramanujan 猜想之间的联系. 我们前面已经说过, 对于群 $\Gamma = SL_2(\boldsymbol{Z})$ 的权 12 尖点模形式空间是一维的, 生成元为

$$\Delta(q) = q\prod_{q \geqslant 1}(1-q^n)^{24} = \sum_{n \geqslant 1}\tau(n)q^n \quad (q = e^{2\pi i\tau}, \tau \in \boldsymbol{H}).$$

Ramanujan 猜想是说：对每个 $n\geqslant1$，

$$|\tau(n)|\leqslant n^{\frac{11}{2}}d(n),d(n)=\sum_{d|n}d.$$

由于 $\Delta(q)$ 是 Hecke 模形式，即有欧拉无穷乘积展开（这最早由 Mordell 于 1917 年发现）

$$\sum_{n\geqslant1}\tau(n)n^{-s}=\prod_{P}(1-\tau(p)p^{-s}+p^{11-2s})^{-1},$$

可知只要对每个素数 p 证明 $|\tau(p)|\leqslant2p^{\frac{11}{2}}$ 即可．而这又相当于证明对每个素数 p，

$$1-\tau(p)U+p^{11}U^2=(1-\alpha(p)U)(1-\beta(p)U),$$

$$|\alpha(p)|=|\beta(T)|=p^{\frac{11}{2}}.$$

如果有 $\boldsymbol{F}_p$ 上的非奇异射影代数簇 X，使得 $1-\tau(p)U+p^{11}U^2$ 是 zeta 函数 $Z_X(U)$ 当中多项式 $P_{11}(U)$ 的因子，由黎曼猜想的模拟便知 $|\alpha(p)|=|\beta(p)|=p^{\frac{11}{2}}$，从而推出 Ramanujan 猜想．在 1965 年—1967 年期间，Eichler，Shimura（志村），Kuga，Ihara 等人构作出了这样的代数簇 X，但是需要把 X 加以紧化．德林给出 X 的紧化方法，从而由黎曼猜想的模拟可以推出 Ramanujan 猜想．

Petersson 把 Ramanujan 猜想作了推广：对于群 $\Gamma_1(N)$ 的每个权 k 的 Hecke 尖点模形式 $f(q)=\sum_{n\geqslant1}a_nq^n$，均有 $|a_p|\leqslant2p^{\frac{k-1}{2}}$（对每个素数 p）．1969 年，德林证明了：由高维韦依猜想可以推出 $k\geqslant2$ 情形的 Petersson 猜想（$k=1$ 情形的 Petersson 猜想是德林和 Serre 于 1974 年证明）．但是问题在于如何证明黎曼猜想的模拟？

1939 年 Rankin 给出了当时对于 $\tau(n)$ 的最好估计 $\tau(n)=O(n^{\frac{29}{5}})$，其方法是研究狄里赫利级数 $\sum_{n\geqslant1}\tau(n)^2n^{-s}$．1970 年朗兰兹指出 Rankin 的方法可用来证明 Ramanujan 猜想，这只需要弄清由 $\Delta(z)$ 给出的无穷多个狄里赫利级数

$$\prod_{p}\prod_{i=0}^{2n}(1-\alpha(p)^{i}\beta(p)^{2n-i}p^{-s})^{-\binom{2n}{i}} \quad (n=1,2,3,\cdots)$$

的极点性状. 德林仔细研究了 Rankin 的 1939 年文章和朗兰兹的上述想法, 发现利用 Grothendieck 的上同调理论和 monodromy 群理论可以得到上述那批狄里赫利级数的极点性状. 更一般地, 他采用 Rankin 方法和 Lefschetz 束的 monodromy 群理论证明了有限域上高维代数簇的黎曼猜想的模拟. 于是他完全证明了高维韦依猜想, 而 Ramanujan-Petersson 猜想是它的推论.

高维韦依猜想是韦依在证明了一维情形之后, 根据高维代数簇的一些计算例子和特征零的复流形几何学与代数拓朴学的启发之下提出的. 25 年以后德林的证明基本上沿着韦依设计的路线, 使用了 Serre 和 Grothendieck 等人在 20 世纪 50 年代和 60 年代建立的庞大的抽象代数几何学, 同时吸取了复几何与模形式理论等方面的新结果和新思想. 但是, 高维韦依猜想本身的叙述形式是非常初等的. 事实上, 这个猜想还可以叙述成更加初等的形式. 设 X 是 $\boldsymbol{F}_q$ 上 d 维非奇异射影代数簇, N_m 表示 X 在 $\boldsymbol{F}_{q^m}$上射影点的个数. 1953 年, S. Lang 和韦依证明了: 若韦依猜想成立, 则对每个 $m\geqslant 1$, 均有

$$|N_m-q^{dm}|\leqslant cq^{m(d-1/2)}, \tag{3.4}$$

其中 C 是依赖于 X 但是与 m 无关的常数, 反过来也可以证明: 若 (3.4) 式对所有 $m\geqslant 1$ 成立, 则也可以推出韦依猜想对 X 成立 (或者更确切地说, 可以推出韦依猜想最核心部分, 即黎曼猜想的模拟对 X 成立). 对于一维情形 (即 X 是 $\boldsymbol{F}_q$ 上非奇异代数曲线), 我们知道由韦依猜想推出 $|N_m-q^m|\leqslant 2gq^{m/2}$ $(m=1,2,3,\cdots)$, 其中 g 是曲线 X 的亏格. 1979 年, Bombieri 只使用黎曼-洛赫定理 (和曲线的伽罗华覆盖思想) 证明了 $|N_m-q^m|\leqslant cq^{m/2}$ $(m\geqslant 1)$, 从而给出一维韦依猜想一个非常初等的证明, 避免了韦依所采用的复杂代数几何机制. 但是对高维情形 $(d\geqslant 2)$, 目前还没有直接证明 (3.4)

式的"初等"方法,所以高维韦依猜想至今没有找到比较初等的证明方法.

为了证明曲线上的猜想,韦依写了三本著作.这些工作的意义不仅在于证明了一维韦依猜想,而更重要的是把代数几何学的研究推进到一个新的时代.正如 Serre 于 1999 年应法国科学院所写的介绍韦依生平事迹的文章中所说的:"韦依写的三百页著作《代数几何基础》很难读懂,而在 20 年后这本书被 Grothendieck 更巨大更难懂的《代数几何原理》所代替."在德林之前,人们对于 Grothendieck 等人发展的抽象代数几何执有不同的看法,不少人怀疑这套庞大的机制究竟有什么用途. Grothendieck 本人在 1970 年以后也"金盆洗手",基本上离开了数学研究职业.在德林使用 Grothendieck 理论证明了高维韦依猜想之后,情形发生了剧然的变化,到了 70 年代末期,这套概型(Scheme)语言和上同调机制已经被许多人(特别是年青数学家)所熟悉和掌握,并已成为研究现代代数几何与数论(算术几何)的通用语言和基本工具.

韦依猜想(现在可以叫作德林-韦依定理)是有限域上非奇异代数簇的理论.人们猜想它和整体域 K 上的代数簇有深刻的联系.以 $K=\boldsymbol{Q}$ 为例,设 X 为定义在 $\boldsymbol{Q}$ 上的非奇异射影代数簇,不妨设定义 X 的代数方程组的系数均属于 $\boldsymbol{Z}$,则存在由有限个素数组成的集合 S,使得对 S 之外的每个素数 p,X 在 p 有好的约化,即该方程组模 p 成为有限域 $\boldsymbol{F}_p$ 上的非奇异代数簇 X_p.于是有韦依 zeta 函数 $Z_{X_p}(U)=\prod\limits_{i=0}^{2d}P_i(X_p,U)^{(-1)^{i+1}}$,其中 d 为 X 的维数.现在对每个 $i(0\leqslant i\leqslant d)$,可以定义整体的 L-函数

$$L(i,X,s)=\prod_p P_i(X_p,q^{-s})^{-1}$$

(当 $p\in S$ 时无穷乘积中的 p-因子采用另外"合理"的定义方式),这叫作 X 的第 i 个 Hasse-Weil L-函数.由韦依猜想可知这个无穷

级数和它展开之后得到的狄里赫利级数 $\Sigma a_n n^{-s}$在 $\mathrm{Re}(s)>1+i/2$ 时收敛，人们猜想

(1) $L(i,X,s)$可以解析延拓成整个复平面上的亚纯函数，并且有函数方程把 $L(i,X,s)$和 $L(i,X,i+1-s)$联系起来.

(2) $L(i,X,s)$在半平面 $Re(s)\leqslant\frac{i+1}{2}$中的所有零点都在直线 $Re(s)=\frac{i+1}{2}$之上. 这也叫作广义黎曼猜想.

从算术几何的观点，$\boldsymbol{Q}$ 上的代数簇对应着由有限域上一个曲线所参数化的一批代数簇，而 Hasse-Weil L-函数对应着这批代数簇的 l-adic 上同调给出的各种 l-adic 表示的 L-函数. 所以 Hasse-Weil L-函数的上述两个猜想应当由有限域上代数簇对于 l-adic 表示的 L-函数的亚纯性质、函数方程和广义黎曼猜想推导出来. 但是在实际上，目前只对于很有限的情形证明了关于 Hasse-Weil L-函数的上述两个猜想. 当 X 是椭圆曲线时，猜想(1)就是著名的谷山-志村-韦依猜想(TSW 猜想). 1956 年，Deuring 对于有复乘的椭圆曲线证明了 TSW 猜想. 在费马猜想的刺激下，怀尔斯对于半稳定的椭圆曲线证明了 TSW 猜想. 到 1999 年，TSW 猜想已被完全解决. 换句话说，猜想(1)对于 $\boldsymbol{Q}$ 上所有椭圆曲线均成立. 1952 年，韦依对于由对角型方程 $a_1x_1^{d_1}+a_2x_2^{d_2}+\cdots+a_nx_n^{d_n}=0$ 定义的超曲面证明了猜想(1). 1958 年，志村对于由模函数参数化的曲线(以及一批重要的代数簇，后人称为 Shimura 簇)证明了猜想(1). 关于猜想(2)目前没有任何重要的结果和突破.

关于有限域上高维代数簇的德林-韦依定理在数论和其他领域有许多深刻的应用. 由它证明了 Ramanujan-Petersson 猜想，是模形式和自守函数理论的重要进展，并且在 1988 年用此可构作性能良好的通信网络(Ramanujan 图，见 § 5.4). 由德林-韦依定理的等价形式 $|N_m-q^{md}|\leqslant c_i q^{m\left(d-\frac{1}{2}\right)}$ 可以对许多类型的指数和给出上界最好的估计，这些估计不仅在解析数论、代数数论和模形式理论

中有重要理论应用，而且在通信中也有诸多实际应用.

1973 年德林证明高维韦依猜想是特征 p(有限域上)的算术几何的巨大进步. 10 年后，特征零(整体域上)的算术几何又得到重大的突破，这就是法廷斯于 1983 年证明了莫代尔猜想.

参考文献

[1] P. Deligne, La conjecture de Weil I, Publ. IHES 43(1974), 273～307.

[2] P. Deligne, La conjecture de Weil II, Seminar at IHES, 1973/1974.

[3] B. Dwork, On the rationality of the zeta function of an algebraic variety, Amer. J. Math. 82(1960), 631～648.

[4] S. Lang and A. Weil, Number of points of varieties in finite fields, Amer. J. Math. 76(1954), 819～827.

[5] R. A. Rankin, Contributions to the theory of Ramanujan's function $\tau(n)$ and similar arithmetical functions II, Proc. Cambridge Phil. Soc. 35(1939).

[6] J.-P. Serre, Faisceaux algebriques coherents, Annals of Math., 61(1955), 197～278.

[7] G. Shimura, Correspondances modulaires et les fonctions ζ de courbes algebriques, J. Math. Soc. Japan, 10(1958), 1～28.

[8] A. Weil, On the Riemann Hypothesis in function fields, Proc. Nat. Acad. Sic. U. S. A. 27(1941), 345～347.

[9] —, Foundation of Algebraic Geometry, Amer, Math. Soc.

Colloq. Pub. ,vo 129,1946.

[10] —,Sur les courbes algebriques et les varietes qui s'en deduisent,Hermann,Paris,1948.

[11] —,Varietes abeliennes et courbes algebriques,Hermann,Paris,1948.

[12] —,Number of solutions of equations in finite fields,Bull. Amer. Math. Soc. 55(1949),497～508.

[13] —,Jacobi sums as Grossencharacter,Trans. Amer. Math. Soc. 73(1952),487～495.

§4.5 法廷斯证明莫代尔猜想(1983 年)

莫代尔(Mordell)猜想是说:有理数域 Q 上定义的任何亏格≥2 的不可约代数曲线均只有有限个有理点.

韦依常常对于人们把一些数学问题称之为"猜想"提出非议.他说:"如果没有经过努力得到大量例子来支持的一个推断,或者对这个推断没有基于理性思考的直觉,这种推断只能是一个数学问题,不能称之为猜想."他还特别以莫代尔猜想为例,说它只是一个数学问题,对它成立或不成立均缺乏足够的依据.莫代尔于 1922 年提出这个猜想时,确实没有太多的例子支持这个猜想,又没有任何反例.其中一个例子是费马曲线 $X^n+Y^n=1$.这个曲线的亏格为 $\frac{(n-1)(n-2)}{2}$,所以当 $n\geqslant 4$ 时,亏格≥3.库默尔证明了当 $3\leqslant n\leqslant 100$ 时,方程 $x^n+y^n=z^n$ 不存在整数解使得 $xyz\neq 0$.由此可知当 $n\geqslant 3$ 时费马曲线 $X^n+Y^n=1$ 只有有理数解 $(X,Y)=(\pm 1,$

0),(0,±1)(当 n 为偶数时),(1,0)和(0,1)(当 n 为奇数时),即莫代尔猜想对于费马曲线 $X^n+Y^n=1(4\leqslant n\leqslant 100)$ 是对的. 反过来说,如果莫代尔猜想成立,那么当 $n\geqslant 4$ 时方程 $X^n+Y^n=1$ 只有有限多有理数解 $(X,Y)=\left(\frac{a}{c},\frac{b}{c}\right)$,其中 $a,b,c\in\mathbf{Z}$. 不妨设 $(a,b,c)=1$,这就推出当 $n\geqslant 4$ 时,方程 $x^n+y^n=z^n$ 只有有限多个本原整数解 $(x,y,z)=(a,b,c)$(本原是指 a,b,c 均为非零整数并且没有大于 1 的公因子). 虽然这与费马猜想还有很大的距离(费马猜想是说 $x^n+y^n=z^n(n\geqslant 3)$ 根本没有本原整数解),但它是对所有 $n\geqslant 4$ 的一个论断,仍可看成是费马猜想的一个重要的进步. 当然,费马曲线只是莫代尔猜想的一个小小的例子. 在猜想提出 60 年之后,对每个非零整数 $a\in\mathbf{Z}$,人们都不知道亏格为 2 的方程 $y^2=x^5+a$ 是否只有有限多个有理数解.

尽管如此,在整整 60 年之后,当时 28 岁的德国年青人法廷斯(Gerd Faltings)于 1983 年证明了莫代尔猜想. 事实上,法廷斯证明了更为一般的结果:对每个代数数域 K,定义于 K 上的亏格 $\geqslant 2$ 代数曲线均只有有限多个 K-点. (对于 K 为任意函数域的情形,类似的猜想由 Manin 和 Grauert 于 1963 和 1965 年分别给出独立的证明.)而且法廷斯同时证明了算术几何中三个重要的猜想:除了莫代尔猜想之外,他还证明了 Tate 猜想和 Shafarevich 猜想. 我们在本节简要地描述一下这三个猜想之间的联系以及法廷斯证明的基本方法和思路.

(一) Shafarevich 于 1962 年提出的猜想为:对于固定的 g,固定的代数数域 K 和有限个 K-位(即 K 的素除子)组成的集合 S,只有有限多个定义于 K 上的亏格 g 代数曲线(同构类),使得这些曲线在 S 之外的每个 K-位上均有好约化.

1968 年,Parshin 对于函数域情形证明了 Shafarevich 猜想. 他同时指出,对于数域的情形,由 Shafarevich 猜想可推出莫代尔猜想. 其推理过程为:假设 C 是亏格 $g\geqslant 2$ 的 K 上曲线,P 为 C 上

一个 K-点，C 在 S 之外的 K-位上均有好的约化．则存在 K 的某个有限扩域 K' 和定义于 K' 上的一条曲线 C'，使得 C' 在 S' 以外的 K'-位上均有好的约化，并且存在有理映射 $C' \to C$，此映射只在点 P 处分歧，而且 K'，S' 和 K' 的亏格 g' 只依赖于 K，S 和 g，而与点 P 无关．Shafarevich 猜想告诉我们：这样的曲线 C' 只有有限多个同构类．如果 C 上有无穷多个 K-点 $P_r(r=1,2,\cdots)$，则在如上得到的曲线 $C'_r(r=1,2,\cdots)$ 当中必有无穷多条同构于一个公共曲线 C''，于是从 C'' 到 C 存在无穷多个不同的有理映射．但是经典黎曼曲面理论告诉我们这是不可能的，从亏格 $\geqslant 2$ 的一个黎曼面到另一个亏格 $\geqslant 2$ 的黎曼面的有理映射只有有限多个．由此矛盾可知，C 上的 K-点只有有限多个．从而由 Shafarevich 猜想推出莫代尔猜想．

（二）上面的 Shafarevich 猜想叫作关于曲线的 Shafarevich 猜想．法廷斯并不是直接证明这个猜想，而是由 Shafarevich 另一个猜想推出它来．这另一个猜想是说：对于给定的 g 和数域 K 以及 K-位的一个有限集合 S，定义在 K 上的 g 维主极化阿贝尔簇只有有限多个在 S 之外均有好的约化．这叫作关于阿贝尔簇的 Shafarevich 猜想．

一个阿贝尔簇即是紧的连通代数簇 V，并且具有有理映射 $V \times V \to V$，使得 V 对于这个映射是阿贝尔群．典型的复阿贝尔簇为 $\mathbf{C}^n/L$，其中 L 是 $\mathbf{C}^n$ 的一个格．一维阿贝尔簇 $\mathbf{C}/L$ 就是椭圆曲线．对于每个定义于数域上的曲线 C，它的全部复数点形成亏格 g 的紧黎曼面．其上的全纯微分 1-形式全体构成 g 维复向量空间 V，对于黎曼面上每条路 p，积分元素 $\int_p$ 属于 V 的对偶空间 V^*（$\simeq \mathbf{C}^g$）．当 p 是闭路时，所有 $\int_p$ 形成 V^* 的一个格．阿贝尔簇 $J=V^*/L$ 叫作是曲线 C 的雅柯比，可以用纯代数方法构作曲线 C 的雅柯比 J，并且 J 也可定义在数域 K 上．根据阿贝尔簇的一个结果（Torelli 定理），J 和它的极化可以完全决定曲线 C．这表明由关于

阿贝尔簇的 Shafarevich 猜想可以推出关于曲线的 Shafarevich 猜想,再由(一)中所述又可推出莫代尔猜想.所以法廷斯的工作实质上是证明关于阿贝尔簇的 Shafarevich 猜想.

(三) 现在谈 Tate 猜想.对于复阿贝尔簇 $A=\mathbf{C}^g/L$,它的子群 $A[n]=\{a\in A\,|\,na=0\}$ 同构于加法群 $\frac{1}{n}L/L\cong(\mathbf{Z}/n\mathbf{Z})^{2g}$.如果 A 是定义在数域 K 上,将 $A[n]$ 中所有点的坐标添加到 K 中,形成 K 的一个扩域 $K^{(n)}$,$K^{(n)}$ 是 K 的代数闭包 K^{ac} 的子域.伽罗华群 $G=\mathrm{Gal}(K^{ac}/K)$ 作用在子群 $A[n]$ 上,给出 G 在 $GL_{2g}(\mathbf{Z}/n\mathbf{Z})$ 上的伽罗华表示.现在取素数 l,则有自然满同态 $A[l^{m+1}]\to A[l^m]$,$a\mapsto la$.于是有极限 $T_l(A)=\varprojlim\limits_{m\to\infty}A[l^m]$,这叫作阿贝尔簇 A 的 Tate 模.由极限过程知 $T_l(A)$ 群同构于 $\varprojlim\limits_{m\to\infty}(\mathbf{Z}/l^m\mathbf{Z})^{2g}=\mathbf{Z}_l^{2g}$,其中 $\mathbf{Z}_l$ 是 ladic 整数环.群 G 通过取极限作用在 $T_l(A)$ 上,从而给出 G 在 $GL_{2g}(\mathbf{Z}_l)$ 中的 l-adic 表示.进而,若 B 是定义在 K 上的另一个阿贝尔簇,则所有从 A 到 B 的群同态形成加法群 $\mathrm{Hom}(A,B)$,而与 G 作用可交换的从 $T_l(A)$ 到 $T_l(B)$ 的群同态形成群 $\mathrm{Hom}_G(T_l(A),T_l(B))$. Tate 猜想是说:

(1) G 到 $T_l(A)$ 上的 l-adic 表示 $G\to GL_{2g}(\mathbf{Z}_l)$ 是半单的.

(2) 有群同构 $\mathrm{Hom}(A,B)\otimes\mathbf{Z}_l\cong\mathrm{Hom}_G(T_l(A),T_l(B))$.

每个同态 $\varphi: A\to B$ 自然诱导出 Tate 模之间的一个同态 $T_l(\varphi)$: $T_l(A)\to T_l(B)$.而猜想(2)本质上相当于说:φ 由 $T_l(\varphi)$ 所决定,即阿贝尔簇之间的同态由它在 Tate 模上的作用所决定,并且 $T_l(A)$ 到 $T_l(B)$ 的每个 G-同态都是由某个 $\varphi: A\to B$ 诱导出来的.

法廷斯首先证明了 Tate 猜想,然后由 Tate 猜想再推出关于阿贝尔簇的 Shafarevich 猜想,由上所述,这也就证明了关于曲线的 Shafarevich 猜想和莫代尔猜想.

法廷斯证明 Tate 猜想和关于阿贝尔簇的 Shafarevich 猜想的方法本质上是费马于三百年前发明的无穷下降法.费马在证明方程 $x^4+y^4=z^4$ 没有正整数解时,他假定 $(x,y,z)=(a,b,c)$ 是一组

正整数解，由此又导出另一组正整数解$(x,y,z)=(a',b',c')$，使得$c'<c$. 但是正整数不能无穷下降，由此导出矛盾. 这个证明本质上利用了正整数的大小概念，并且对每个正实数c，不超过c的正整数只有有限多个. 为了证明关于阿贝尔簇的Shafarevich猜想，即定义在数域K上的g维主极化阿贝尔簇只有有限多个，使得它们在S之外的每个K-位上都有好的约化，法廷斯对每个这样的阿贝尔簇A定义一个衡量大小的量$h(A)$，叫作A的**高度**. 这个高度的定义比较复杂，但是它可以转化成另一个比较容易叙述的高度概念. Siegel把K上的所有g维主极化阿贝尔簇作成一个新的射影代数簇n_g，叫作参量空间(moduli space). 也就是说，K上的每个g维主极化阿贝尔簇A是n_g中的一个点x. 由于n_g是射影代数簇，它可嵌到K上的某个射影空间$\boldsymbol{P}^n(K)$中. 我们可以在射影空间中定义一个高度概念. 以$K=\boldsymbol{Q}$为例，n维射影空间$\boldsymbol{P}^n(\boldsymbol{Q})$的每个点可以惟一表示成$x=(x_0,x_1,\cdots,x_n)$，其中$x_0,x_1,\cdots,x_n$是不全为零的整数并且没有大于1的公因子. 点$x$的高度定义为$h(x)=\max\limits_{0\leq i\leq n}\{\log|x_i|\}$. 不难看出，对于每个正实数$c$，$\boldsymbol{P}^n(\boldsymbol{Q})$中满足$h(x)\leq c$的射影点$x$只有有限多个. 对于一般的代数数域$K$，射影空间$\boldsymbol{P}^n(K)$中点$x$的高度$h(x)$也可以定义，并且有同样的性质：高度有界的点只有有限多个. 如果阿贝尔簇A对应于点x，则法廷斯证明两个高度$h(A)$和$h(x)$本质上是一回事.

现在，K上每个g维主极化阿贝尔簇A对应参量空间n_g的一个点，从而对应射影空间$\boldsymbol{P}^n(K)$中一个点. 而其中在S外的K-位均有好约化的阿贝尔簇形成n_g的一个子集合n'_g，它也是射影空间$\boldsymbol{P}^n(K)$中的一个子集合. 利用Tate猜想可以证明$\boldsymbol{P}^n(K)$中子集合n'_g所有点的高度是有界的(证明中还使用了表示论等一系列结果，甚至还利用了上节讲述的德林-韦依定理)，所有n'_g是有限集合，从而只有有限多个阿贝尔簇在S之外的K-位均有好的约化，这就由Tate猜想证明了关于阿贝尔簇的Shafarevich猜想.

法廷斯证明Tate猜想也是使用了上述的“有界高度原则”，并

且在技术细节中也使用了 Zarhin 等人的其他重要结果.

以上我们简要地介绍了法廷斯证明包括莫代尔猜想在内的三个重要猜想的证明思想和程序.从介绍中我们看到,法廷斯使用了前苏联代数几何学派(Shafarevich,Parshin,Zarhin,Manin)等人的工作和美国哈佛代数几何学派(Tate 等人)的工作.事实上,这两个学派长时期所从事的研究都把证明莫代尔猜想作为重要目标之一.正如哈佛学派另一位重要成员 B. Mazur 所作的评论:"法廷斯所作的事情,就是他意识到把这两方面代数几何成果合在一起就可以证明莫代尔猜想,而我们每个人都没有看出这一点."

法廷斯的这项工作获得 1986 年 Fields 奖.这项工作把数论和算术几何的研究推进到一个新的时期.它不但继德林之后再一次显示出 Grothendieck 代数几何的威力,而且使人们意识到,对于整体域上的许多重要的数论问题,不仅需要利用局部域工具,p-adic 工具和表示理论,而且还需要复微分几何.这就引起许多优秀的微分几何学家对数论的兴趣.

以费马猜想为例.我们在前面说过,由莫代尔猜想(现在可以叫作法廷斯定理)可以推出,对每个 $n\geqslant3$,方程 $x^n+y^n=z^n$ 只有有限多个本原整数解.虽然这离完全证明费马猜想还有相当的距离,但是这重新激发起人们研究费马猜想的兴趣.自从 1847 年库默尔关于费马猜想的工作之后,关于费马猜想的研究主要有两个方面.一个方面是给出 n 的各种充分性条件,使得方程 $x^n+y^n=z^n$ 没有正整数解.但是在 1983 年以前,人们都不能证明存在无限多个素数 p,使得 $x^p+y^p=z^p$ 没有正整数解.第二个方面是利用这些充分性条件再加上计算机计算,例如在 1977 年证明了当 $p<125000$ 时 $x^p+y^p=z^p$ 没有正整数解.但是这对完全解决费马猜想没有任何启示.概括地说,在库默尔之后的一百多年里,尽管有许多关于费马猜想的结果,但是在思想和方法上没有本质性突破.人们对完全解决费马猜想看不出任何希望,以致于有不少著名数学家认为解

决费马猜想是 21 世纪的事情. 在法廷斯工作之后,人们感到采用几何方法有可能使费马猜想取得重大突破. 到了 20 世纪 80 年代后期对于通往费马猜想的途径人们设计出许多方案,其中最主要的有以下三个:

(1) 利用算术曲面的博戈莫洛夫-米亚奥卡-邱成桐不等式. 由这个不等式可推出对于充分大的素数 p,方程 $x^p+y^p=z^p$ 没有正整数解. 但是上述不等式至今未被证明.

(2) 利用曲线上关于整点高度理论的 Vojta 猜想. 由它可推出对充分大的正整数 n,方程 $x^n+y^n=z^n$ 没有正整数解. 但是 Vojta猜想至今也未被证明.

(3) 发现由椭圆曲线的谷山-志村-韦依猜想可以推出整个费马猜想的正确性.

这三条路线都属于算术几何. 几何学家们热衷于前两条路线,第三条路线更偏重于数论,并且多数学者认为要证明谷山-志村-韦依猜想是非常困难的. 后来的历史表明,怀尔斯从 1986 年起默默无闻地埋头苦干了八年,终于走通了第三条路线,于 1994 年证明了费马猜想.

参 考 文 献

[1] G. Faltings, Endlichkeitsatze fur abelsche Varietaten uber Zahlkörpern, Invent. Math. 73(1983), 349～366.

[2] P. Deligne, Preuves des conjectures de Tate et de Shafarevich, (d'après G. Faltings)Bourbaki n. 616(1983～1984).

[3] Y. Manin, Rational points of algebraic curves over functional fields. Izv. Akad. Nauk SSSR, Ser. Mat. 27(1963), 1395～

1440.

[4] L. J. Mordell, On the rational solutions of the indeterminate equations of the third and fourth degrees, Proc. Camb. Philos. Soc. 21(1922), 179～192.

[5] H. Grauert, Mordell's Vermutung uber rationale Punkte auf algebraischen Kurven und Funktionen körper, Publ. Math. IHES vol 25(1965), 363～381.

[6] L. Szpiro, Le conjecture de Mordell(d'apres G. Faltings), Sem. Bourbaki, Exp. 619(1983).

[7] A. N. Parshin, Algebraic curves over function fields I, Izv. Math. USSR. 2(1968), 1145～1170.

[8] J. Tate, Endomorphisms of abelian varieties over finite fields, Inv. Math. 2(1966), 134～144.

[9] J. G. Zarhin, Isogenies of abelian varieties over fields of finite(haracteristics, Math. USSR Sbornik 24(1974), 451～461.

[10] S. Bloch, The proof of the Mordell conjecture, Mathematical Intelligencer, 6(1984), 41～47.

[11] G. Faltings, G. Wüstholtz et al., 《Rational Points》, Seminar Bonn/Wuppertal 1983/1984, Third Enlarged Edition, 1992.(此书是 1983/1984 年关于 Faltings 工作研讨会文集 1992 年的第三版(扩大). 书中由专家们(包括 Faltings 本人)对各种相关知识作了系统的介绍.)

[12] L. Szpiro, Seminaire sur les pinceaux arithmétiques: La conjecture de Mordell, Asterisque 127(1985).

[13] I. R. Shafarevich, Algebraic number fields, Proc. Internat. Congr. Math. Stockholm 1962, Institute Mittag-Liffler, 1963. pp. 163～176.



安德鲁·怀尔斯

§4.6 怀尔斯证明费马猜想(1994年)

德国黑森林州中部有一个景色宜人的幽静小镇,名叫 Oberwalfach.多年来,这里成为世界各地数学家的"旅游区".每年举行

几十个为期一周的数学高级研讨会，由数学各领域的世界顶尖高手主持，研讨热门的数学问题，交流数学成果和思想. 这里没有卡拉 OK，私人房间里没有电视机，有的是黑板、图书、计算机房和供数学家们交谈的咖啡室. 星期六，参加上一个会议的数学家们满载而归，而第二天，下一个会议的参加者前来报到.

1984 年秋，一群优秀的数论学家在这里聚会，研究和讨论关于椭圆曲线的各种突破性工作. 其中一位演说者——德国数论学家弗雷(Gerhard Frey)的报告使所有听众振奋和惊喜. 众所周知，如果对每个奇素数 p，方程 $x^p+y^p=z^p$ 均没有整数解使得 $xyz\neq 0$，那末整个费马猜想便被证明了. 弗雷说，假如此方程有一组整数解 $(x,y,z)=(a,b,c)$，$abc\neq 0$，然后他在黑板上写出一条椭圆曲线

$$y^2=x(x-a^p)(x+b^p)$$

(这条曲线后来便称为**弗雷曲线**). 然后他推导出这条曲线有一系列奇特的性质，最后导致这条曲线不满足谷山—志村—韦依猜想(TSW 猜想)！也就是说，若费马猜想不成立，则关于椭圆曲线著名的 TSW 猜想也不能成立，即由 TSW 猜想可推出费马猜想！

除了弗雷本人，几乎在场的所有听众都发现讲演中有一个明显的逻辑错误，但这并没有减弱人们的兴趣，每个人都希望自己是第一个能够补救这个缺欠的人. 几个月过去了，电子邮件上没有传出这个缺欠被弥补的任何消息，人们开始意识到这件事不像看上去的那么容易. 一直到 1986 年夏天，世界数学家大会在美国加州大学伯克利分校举行. 弗雷讲演的两位听众—伯克利的里贝特(K. Ribet)和哈佛大学的梅祖尔(B. Mazur)在咖啡店里聊天，里贝特向梅祖尔讲解他在近两年来试图完善弗雷讲演所采取的策略和遇到的困难. 突然，梅祖尔说："难道你还不明白？你已经完成了它！你还需要做的只是加上一些 M-结构，然后再做一遍你的论证就行了."事后里贝特回忆说："我怎么会不明白这一点，我完全惊呆了，因为我从未想到添加额外的 M-结构，听上去如此简单.""我高兴得像上了天似地回到住所……. 我完全被迷住了，

我坐下来，开始在笔记本上飞速地写起来．大约两个小时后，我已经写完了一切．……国际数学家大会当然有成千的数学家参加，我有点随便地对几个人提到我已经证明了由 TSW 猜想可推出费马猜想．这消息象野火般传了开来……”

TSW 猜想的通常形式是说：$\boldsymbol{Q}$ 上每条椭圆曲线的 L-函数都是某个权 2Hecke 模形式的 L-函数．我们已经知道这是朗兰兹纲领的一个源头，所以和群 Gal $(\boldsymbol{Q}^{ac}/\boldsymbol{Q})$ 的伽罗华表示有密切联系．具体说来，$\boldsymbol{Q}$ 上的每条椭圆曲线 E 的全部复点 $E(\boldsymbol{C})$解析同构于 $\boldsymbol{C}/L$，其中 $L=w_1\boldsymbol{Z}\oplus w_2\boldsymbol{Z}$ 是 $\boldsymbol{C}$ 中一个秩 2 的格．以 $E[p^n]$ 表示 $E(\boldsymbol{C})$中 p^n 阶元素全体，它同构于$\frac{1}{p^n}L/L\cong(\boldsymbol{Z}/p^n\boldsymbol{Z})^2$（其中 p 是任何素数）．当 $n\to\infty$时，我们便知道 $E(\boldsymbol{C})$中子群

$$T(p)=T(E,p)=\{a\in E(\boldsymbol{C})\mid \text{存在 } n\geqslant 0 \text{ 使得}[p^n]a=0\}$$

同构于$\lim\limits_{n\to\infty}(\boldsymbol{Z}/p^n\boldsymbol{Z})^2=\boldsymbol{Z}_p^2$（其中 $\boldsymbol{Z}_p$ 是 $\boldsymbol{Q}_p$ 的 p-adic 整数环）．$T(p)$ 叫作 E 的 Tate 模．群 $G=$Gal $(\boldsymbol{Q}^{ac}/\boldsymbol{Q})$ 在 $T(p)$上的自然作用给出 G 到 $GL_2(\boldsymbol{Z}_p)$上的一个伽罗华表示 ρ_p．模 p 之后还得到 G 在 $GL_2(\boldsymbol{F}_p)$上的伽罗华表示 $\bar{\rho}_p$．假如 TSW 猜想成立，那么这些表示应当是由模形式构作的自守表示．人们已经知道这些自守表示（对每个素数 p）有许多特殊的性质，具有这些特殊性质的表示叫作**模表示**．1987 年，Serre 曾提出一个猜想，Serre 猜想的一部分是说：对每个有限域 F，Gal $(\boldsymbol{Q}^{ac}/\boldsymbol{Q})$ 在 $GL_2(F)$上的每个绝对不可约守表示都是模表示．而弗雷的讲演的要旨是说，如果 $x^p+y^p=z^p$ 有整数解$(x,y,z)=(A,B,C)$，$ABC\neq 0$，则他那条椭圆曲线 $Y^2=X(X-A^p)(X+B^p)$给出的 Gal $(\boldsymbol{Q}^{ac}/\boldsymbol{Q})$ 的伽罗华表示 $\bar{\rho}$ 不是模表示．如果 TSW 猜想成立，则 $\bar{\rho}$ 应是自守表示．又若 Serre 猜想成立，则 $\bar{\rho}$ 应是模表示，这就导出矛盾．所以弗雷演讲真正证明的，是由 TSW 猜想（椭圆曲线和模形式的联系）加上 Serre 猜想（伽罗华表示为模表示）推出了费马猜想．当然，这是里贝特在弗雷讲演之后才弄清楚的．里贝特所作的事情是证明了 Serre

猜想，于是完全确认了费雷的论断：由 TSW 猜想便可推出费马猜想．上述框架也使我们能够理解；1993 年怀尔斯第一次公开宣布他证明费马猜想时，他的讲演标题是“模形式，椭圆曲线和伽罗华表示”．他只字不提费马猜想，他讲演的最后一句话是：“这样，我就对所有半稳定(semistabe)的椭圆曲线证明了 TSW 猜想．”在场的所有专家们那时都已知道，弗雷曲线是半稳定的，所以这就证明了费马猜想．

这样，由于里贝特的工作人们已经确信：由 TSW 猜想可推出费马猜想，甚至只要证明对弗雷曲线证明 TSW 猜想成立就可以了．但是这个事情是艰巨的，甚至连里贝特也持悲观的态度：“绝大多数人相信谷山一韦依猜想是完全无法接近的，我是其中的一个．我没有真的费神去试图证明它，我甚至没有想到过要去试一下．怀尔斯大概是地球上敢大胆梦想可以实际上证明这个猜想的极少数几个人之一．”（里贝特和一些其他数学家在讲演和文章中常常在叙述 TSW 猜想时忽略志村的名字．在受到“抗议”之后，里贝特公开表示要在他介绍怀尔斯工作的文章中补上志村的名字．）

怀尔斯是在 1986 年夏末（伯克利世界数学家大会开后不久）在一个朋友家里饮冰茶时得知里贝特的工作的“我感到极大的震动，我记得那个时刻，那个改变我生命历程的时刻，因为这意味着为了证明费马大定理，我必须做的一切就是证明谷山—志村猜想．它意味着我童年的梦想现在成了体面的值得去做的事．我懂得我决不能让它溜走．我十分清楚我应该回家去研究谷山—志村猜想．”

怀尔斯（Andre Wiles）于 1953 年 6 月 23 日生于英国剑桥，1974 年在剑桥大学莫尔顿学院毕业，1977 年在克莱尔学院获博士学位．1977 年～1980 年，在美国哈佛大学任助理教授，1981 年在普林斯顿高等研究院任研究员，1982 年任普林斯顿大学教授，1984 年任普林斯顿大学欧仁·希金斯讲座教授．在他十几岁的时

候，从一位中学老师给他的数论书中知道了费马猜想，并花了一年的时间充满热情地研究它. 开始时他怀着一种单纯而有抱负的自信：20 世纪的中学生所懂得的数学和 17 世纪的天才费马一样多，或许由于他的天真会使他碰巧找到一个世故的学者未曾注意到的证明，这就是他“童年的梦想”. 一年后的失败使他改变了策略，决定要认真研究前人的方法，从更高明数学家的错误中学到一些有用的东西，“设法理解前人的工作”.

到 1986 年再次唤起他童年的梦想时，怀尔斯已从一个天真的中学生成为一个成熟而杰出的年青数学家，我们在前面已经提到，他和他的老师 J. Coates，在 1977 年对于椭圆曲线的 BSD 猜想做了突破性的工作，他与梅祖尔于 1984 年证明了 $\boldsymbol{Q}$ 上的岩泽主猜想，他在 1978 年对于高次互反律做了出色的工作. 这些工作使他被国际数学界公认为在数论特别是椭圆曲线和模形式方面为数不多的年轻学者之一. 他这时的自信已建立在坚实的研究基础之上. 他又深知，要证明 TSW 猜想是非常艰巨的任务，但他已下定决心. “当然，已经很多年了，谷山—志村猜想一直没有解决. 没有人对怎样处理它有任何想法，但至少它属于数学的主流. ……我不认为我在浪费自己的时间. 这样，吸引了我一生的费马传奇故事现在和一个专业上有用的问题结合起来了.”

怀尔斯放弃了所有与证明 TSW 猜想无直接关系的研究工作，不再参加学术会议和报告会. 除了在普林斯顿教书，指导研究生和参加必要的讨论班之外，回避一切分心的事情，回家躲进他顶楼的书房. 开始他面壁七年的艰苦历程. 这期间他取得了一些突破，但在未完成证明之前，他不与任何人讨论，也不发表任何部分结果. 为了不引起人们的怀疑，他每隔 6 个月发表一篇小论文. 这样做并不是想增加文章的篇数，而是让同事相信他仍在继续他平常的研究，掩人耳目. 结果，在这七年中间，甚至关系密切的同事也没留意到他的潜心所至. 他的老师 Coates 对此也毫不知情：“我记得在许多场合对他讲，与费马猜想的这种联系非常

好，但要证明谷山—志村猜想仍然是毫无希望的. 他当时只是对我笑笑.”里贝特也同样不知怀尔斯暗中进行的工作.“这大概是我知道的仅有的例子，一个人进行了这么长时间的研究而不公开他在做什么，也不谈论他正在取得的进展. 在我的经历中，这是前所未闻的……在我们这个团体中，人们总是分享他们的想法. ……如果你把自己与此隔绝起来，那么从心理学观点来看你是在做非常古怪的事情.”但是怀尔斯有自己的解释：“我意识到与费马大定理有关的任何事情都会引起太多的兴趣. 除非你的专心不被他人分散，否则你不可能很多年都使自己精力集中，而这一点会因旁观者太多而做不到.”当然，这样做的动机不排除对荣誉的追求.（作为众多流言当中的一个，是说他在剑桥大学宣布结果的那一天是他 40 岁的生日，而 Fields 奖只给不超过 40 岁的数学家.）

著名科学记者西蒙·辛格在《费马大定理：一个困惑了世间智者 358 年的谜》一书中对于怀尔斯的 7 年艰辛作了生动的描绘. 从数学角度，他一直试图解决伽罗华表示方面的问题. 但是在 1988 年 3 月 8 日，《华盛顿邮报》和《纽约时报》宣称东京大学 38 岁的宫岗洋一（Yoichi Miyaoka）用几何方法证明了费马猜想时，怀尔斯大吃一惊. 那正是继法廷斯证明莫代尔猜想之后，几何学家们对着费马猜想蜂拥而上的时候，但正是法廷斯这位对几何与数论均在行的数学家指出宫岗的严重漏洞，使怀尔斯松了一口气，仍旧沿自己的道路前进. 作为“宫岗旋风”的一个余波，纽约第八街地铁车站里，有人模仿费马在墙上写下四行美式的幽默：

$x^n+y^n=z^n$ 没有解，
对此我已发现一个真正美妙的证明，
可惜我现在没时间写出来，
因为我要坐的地铁车正开过来.

1990 年，怀尔斯开始考查椭圆曲线的岩泽理论，试图寻找新的突破口. 5 年当中他作了两次父亲，“我放松一下情绪的唯一方

式是和我的孩子们在一起."到了1991年夏天,怀尔斯在当了5年"隐士"之后,感到需要重返数学界,了解数学最新成果,他参加了在波士顿举行的椭圆曲线会议. 正是在这个会上,他的老师Coates告诉他,一个名叫Matheus Flach的学生用Kolyvagin(前苏联年青数论学家)的方法研究椭圆曲线. 回到普林斯顿大学之后,怀尔斯用Kolyvagin-Flach方法对于某些椭圆曲线解决了他的伽罗华表示问题. 但为了达到最终目的,还需发展这种方法. 到了1993年1月,怀尔斯"决定有必要向一个人吐露秘密,而他应该是一位我正在使用的那一类几何方法的专家. 我需要非常小心地挑选这个我要告知秘密的人,因为他必须保守住秘密. 我选择了尼克·凯兹(N. Katz)."

凯兹是怀尔斯在普林斯顿大学数学系的同事. 当怀尔斯向凯兹吐露了秘密并希望帮助他一起检查和讨论极为专门的技术细节之后,凯兹的建议也是颇具有喜剧性的. 怀尔斯向研究生开设了以"椭圆曲线的计算"为题的讨论班,凯兹是听众之一. 一开始怀尔斯就讲述这6年来他所进行的专门计算和证明,研究生们谁也不知怀尔斯到底要干什么,而外人却以为这是在给研究生讲课. 研究生一个个地消失,几个星期之后,听众只剩下凯兹一人.

1993年5月,怀尔斯吸收了梅祖尔文章中的思想之后,相信他解决了TSW猜想的一部分并导致证明了费马猜想."我仍然需要再核对一下证明,但是6月末在剑桥有一个会议要举行,我想这也许是宣布这个证明的好地方:它是我的故乡,我曾在那里念研究生."

剑桥大学牛顿数学所的这次会议名为"L-函数和算术". 会议组织者之一Coates应怀尔斯请求破例地为怀尔斯安排6月21、22、23日上午的三次演讲"'Andry,你究竟证明了什么?我们要不要告诉新闻界?',他只是微微地摇了一下头,依然紧闭双唇. 他确实在为高度戏剧性的场面作准备. 然后就是6月23日上午那场激动人心的局面. 在报告之前怀尔斯只把证明的复印件给了一个

人一梅祖尔，但是梅祖尔对这个演讲仍然感到惊讶，“我从未见过如此辉煌的演讲，充满了如此奇妙的思想，具有如此戏剧性的紧张，准备得如此之好.”里贝特后来回忆道，在怀尔斯演讲之后，“下一位报告人是一个名叫里贝特的人，就是鄙人. 我作了演讲，人们也作了笔记，鼓了掌. 可是在场的每个人，包括我自己，对我在演讲中讲了些什么都没有丝毫印象.”

《卫报》和《世界报》均在头版报道此事，这是自1988年宫岗宣布证明费马猜想以来，数学家又一次占据头条新闻. 唯一的差别是这次的报道量比上次多达两倍，并且没有人对怀尔斯的证明表示怀疑. 对于专业数学家来说，解决TSW猜想是比证明费马猜想更巨大的成就，因为前者对数论有极大的影响，可是新闻报道则主要在说费马猜想，偶而有记者顺便提一下TSW. 一夜之间怀尔斯成了世界著名人士，《People》杂志将他和戴安娜王妃一起列为“本年度25位最具有魅力者”之一，一家国际制衣大公司请这位温文而雅的天才为他们的新系列男装做广告. 而在数学界，报道这个消息的电子邮件几天之内就传到世界各地，正在暑假期间的各大学数学教授和学生们纷纷聚集起来，听会议参加者“传达”怀尔斯的报告.

在一阵热烈而兴奋的情绪过去之后，数学家们并没有忘记对怀尔斯的证明进行严格的审查. 文章寄到《Inventions Mathematicae》杂志，编辑梅祖尔作了特别的决定，挑选了六位审稿人，200页的证明被分成六章，每位审稿人只负责其中的一章. 其中第三章由凯兹审查，8月份凯兹发现一个问题，怀尔斯认为容易补救，但到了9月份，逐渐意识到怀尔斯采用的Kolyvagin-Flach方法仍有障碍. 这个时候，怀尔斯面临他最大的考验，甚至连审稿人都感到了保守秘密带来的压力. 另一个审稿人里贝特说，“我开始被人称为‘费马信息咨询所’，……人们开始变得不耐烦了，开始问一些棘手的问题.”“他们就是想知道后来发生了什么. 以后情况变得更糟，因为慢慢地怀疑的阴影集中到证明本身上了. 人们不

断告诉我这些谣传，说在第三章有缺陷. 他们问我知道些什么，我真的不知道如何回答才好.”

经过一个凄凉的秋天，在各种传闻和情绪的巨大压力下，1993年 12 月 4 日怀尔斯向数学界发了一个电子邮件："鉴于存在着对我的关于谷山—志村猜想和费马大定理工作状况的种种推测，我对情况作一简短说明. 在检验过程中发现许多问题，大部分已经解决，但是有一个特别的问题我还没有解决. ……我相信在不远的将来我能够使用我在剑桥演讲中解释过的想法完成它.”

在最黑暗的时期，怀尔斯向他在普林斯顿大学数学系另一个好友 Peter Sarnak 说，他准备公开承认失败. Sarnak 向他暗示，他应当寻找一个他信赖的人，熟悉 Kolyvagin-Flach 方法的专家经常在一起进行技术细节的讨论，鼓励他用这种方式再试下去. 对这件事又作了长时间考虑之后，怀尔斯决定邀请剑桥大学的一位讲师理查德·泰勒（Richard Taylor）到普林斯顿和他一起工作. 泰勒是怀尔斯以前的学生、怀尔斯文章审稿人之一. 泰勒怀着一种使命感来到普林斯顿，但一直到 1994 年夏天仍没有新的突破.

这年 8 月，世界数学家大会在瑞士苏夷士召开. 怀尔斯已与 Fields 奖无缘，但是他仍被邀请在闭幕式上做最后一个大会报告. 在报告中他坦诚地介绍自己对 TSW 猜想已经证明到什么程度，为了得到费马猜想还需做哪些工作. 这些工作何时能够完成，他说他自己也不知道. 整个大厅以极为热烈的掌声回答他的演讲，这给他很大的安慰，数学界认可他为数论提供了一大套新的技术和策略，他的新结果可用来证明别的定理，他开始感到失败并不耻辱，准备适应受到打击后的境迁.

可是就在会后不久，9 月 19 日，一个星期一的早晨，他准备再一次了解失败的原因时，突然冒出了一个想法：将岩泽理论和 Kolyvagin-Flach 方法结合在一起!”我无法理解我怎么没有发现它，足足有 20 多分钟我呆望着它不敢相信. 然后我到系里转了一圈，又回到桌子旁试图看看是否真的如此. 但情况确实这样. ……

第二天早晨我又一次作了核对，到 11 点时我完全放心了.”简言之，证明最后归结为一个纯代数问题：关于 Hecke 环的完全交性质. 这最后关口是他与泰勒一起共同合作闯过的. 于是，1994 年 10 月 25 日，两篇文章一起寄到国际最权威数学刊物《Annals of Mathematics》一份是怀尔斯署名的《模椭圆曲线和费马大定理》，另一份是他与泰勒合作的《某些 Hecke 代数的环论性质》. 在法廷斯对其中部分论证作了重大简化之后，文章于 1995 年正式发表. 至此，费马猜想正式宣告得到解决. 1996 年，怀尔斯和朗兰兹共同获得沃尔夫奖. 这个奖通常是给予一生中为世界数学做了突出贡献的长者，怀尔斯是第一位获此殊荣的四十多岁的年轻数学家.

在怀尔斯证明费马猜想之后的近 4 年里，代数数论仍有令人瞩目的理论进步(TSW 猜想已完全解决)，但我们宁愿以这个充满艰辛而又美丽动人的故事作为本章的结尾.

参 考 文 献

[1] 《Elliptic Curves, Modular Forms and the Fermat's Last Thorem》, (Second Edition) J. Coates and S. -T. Yau (editors), International Press, 1997.

[2] 《Modular Forms and Fermat's Last Theorem》, G. Cornell, J. H. Silverman, G. Stevens Editors., Springer-Verlag, 1997.

[3] R.L. Taylor and A. Wiles, Ring theoretic properties of certain Hecke algebras, Annals of Math. 141 (1995), 553～572.

[4] A. Wiles, Modular elliptic curves and Fermat's Last Theorem, Annals of Math. 141 (1995), 443～551.

[5] S. Singh, 《Fermat's Last Theorem, The story of a riddle that

confounded the world's great minds for 358 years》Fourth Estate Limited, London, 1997.
(中译本:《费马大定理: 一个困惑了世间智者358年的谜》, 薛密译, 上海译文出版社, 1998.)
[6] 胡作玄:《350年历程: 从费尔马到维尔斯》, 新视野丛书(第1辑), 山东教育出版社, 1996.

第五章　代数数论的应用 (1960—　)

代数数论在 20 世纪取得了辉煌的理论成就. 与此同时，特别是在 20 世纪的后半期，数论在各种实际领域得到广泛和重要的应用，其主要原因是 20 世纪 60 年代以来数字计算机技术和数字通信技术的巨大进步. 在此之前，数学主要应用于力学、物理学等方面，以付立叶分析为代表的分析学用来描述事物连续现象，数理统计学用来描述事物的统计特性和随机现象，成为应用数学的主流. 数字计算机和数字通信技术的出现和飞速发展极大地改变了应用数学的面貌. 这些领域提出了大量的数学课题，使描述事物离散现象的离散数学在坚实的应用根基之上得到很大的促进和发展. 组合数学和图论在过去被认为是雕虫小技和数学游戏（36 军官问题，七桥问题，棋盘跳马问题，哈密顿周游世界问题…）而不能登大雅之堂，现在已逐渐发展成整个数学的一个重要领域. 20 世纪的代数数论在理论研究中与几何、分析数学的融合，体现出人类对世界事物连续性和离散性内在联系的认识的深化. 而计算机科学和信息科学技术的发展使数论在研究风格和目标上也发生了很大变化. 一方面，数论理论研究中使用计算机，使数论本身更加成为实验性学科，另一方面是数论在实验和技术学科的广泛应用，特别是数论近年来取得的重大理论成果在实际中得到应用，使通信等领域得到体制上的突破和进步，被人们称为“令人振奋”和“不可预测”的事件. 目前在世界范围内已有许多优秀的数论学家从事以应用为背景的研究工作，与计算机科学和信息工程专家一起，形成“应用数论”或“计算数论”的强大的队伍. 这个队伍的工作不仅为数论研究本身注入了新的风格和活力，而且对应用和技术进步予以重大的影响.

我们在本章中首先介绍在代数数论研究中从事计算和实验性工作的进展概况. 然后介绍代数数论的重大理论成果在信息科学中三个方面的应用. 关于数论的更广泛应用可参见:

M. R. Schroeder, Number Theory in Science and Communication. with Applications in Cryptography, Physics, Biology, Digital Information and Computing, Springer-Verlag, 1984.

§5.1 计算代数数论

许多数论学家都善长计算. 欧拉对无穷级数进行大量计算,得出 $\zeta(2)=\sum_{n\geqslant 1}n^{-2}=\frac{\pi^2}{6}$,并且给出黎曼 zeta 函数 $\zeta(s)$ 的函数方程雏型. 高斯在 19 岁时算出了 17 次本原单位根的确切值. 库默尔计算出 $\mathbf{Z}[\zeta_p]$($p\leqslant 19$)的类数为 1,而 $\mathbf{Z}[\zeta_{23}]$ 的类数为 3. 这些都是在当时条件下手算的精品. 随着计算机硬件和软件的发展和应用领域需求的推动,近 30 年数论计算工作特别是各种算法改进的研究得到了飞速的发展. 以代数数论为例,可以比较如下两份材料:

[1] H. Zimmer, Computational Problems, Methods and Results in Algebraic Number Theory, LN in Math. 262, Springer-Verlag, 1972.

[2] H. Cohen, A Course in Computational Algebraic Number Theory, Graduate Texts in Math. 138, Springer-Verlag, 1996.

前者介绍了 1971 年以前代数数论在计算问题中的方法和成果,后者是"计算代数数论"课程的研究生教材. 时隔 20 余年,在研

究问题的拓展、算法的改进和计算成果的深入等方面都发生了巨大的变化.

计算代数数论研究的推动力量大体有以下三个方面：

1. 从大量计算数据中探索规律，提出和修正猜想.

费马计算出 $F_n=2^{2^n}+1(n=0,1,2,3,4)$ 均为素数(费马素数)，提出猜想：对所有正整数 n，F_n 均为素数. 一百年后欧拉于 1732 年利用同余式演算巧妙地发现 641 是 F_5 的一个素因子，推翻了这个猜想. 到目前为止还没有发现新的费马素数. 1880 年 Landry 把 F_6 分解成两个素数之积. 又过了 90 年，Morrison 和 Brillhart 发明了大数分解的新算法——连分数算法，于 1970 年分解出 $F_7=P_{17}\cdot P_{22}$，其中 P_t 表示一个 t 位的素数. 10 年后 Brent 和 Pollard 又发明了新的算法，于 1980 年分解出 $F_8=P_{16}\cdot P_{62}$. 1988 年，Brent 用椭圆曲线算法分解出 F_{11}，它是五个素数的乘积，这五个素数的位数分别为 6,6,21,22 和 564. 1990 年，Lenstra 等人采用数域筛法，动用了世界各地 700 多个工作站进行并行计算，分解出 $F_9=P_7\cdot P_{49}\cdot P_{99}$. 1995 年 Brent 用椭圆曲线算法分解了 $F_{10}=P_8\cdot P_{10}\cdot P_{40}\cdot P_{252}$. 目前已知 F_{12}是 7 个素数的乘积，其中 5 个素因子已经求出，剩下的一个 1187 位数应当是两个素数的乘积，但是还没有把它分解出来. 是否有无穷多费马素数？是否有无穷多费马数 F_n 是合成数？目前都不知道. 从目前的数据提任何猜想都为时太早.

另一个例子是对素数分布的研究. 对每个实数 x，以 $\pi(x)$表示不超过 x 的素数个数. 公元前 3 世纪古希腊数学家证明了素数有无穷多个，即 $\lim\limits_{x\to+\infty}\pi(x)=+\infty$. 勒让得于 1798 年给出计算公式 $\pi(N)=\pi(\sqrt{N})-1+\sum\limits_{d|N}\mu(d)\left[\dfrac{N}{d}\right]$，其中 N 为正整数，$\mu(d)$为 Möbius 函数. 由这个公式可以推出 $\lim\limits_{x\to\infty}\dfrac{\pi(x)}{x}=0$，即素数在正整数集合中稀疏得“不成比例”. 勒让得根据 $\pi(x)$的计算数据提出猜想

$$\pi(x)=\frac{x}{\log x-A(x)},\ \lim_{x\to\infty}A(x)=1.08366\cdots$$

但是1850年俄国数学家切比雪夫推翻了这个猜想.他证明了若 $\lim\limits_{x\to\infty}A(x)$ 存在,它一定为1.他同时证明了:若极限 $\lim\pi(x)\Big/\dfrac{x}{\log x}$ 存在,则它一定为1.1792年,15岁的高斯利用他的计算结果猜想当 $x\to+\infty$ 时,$\pi(x)\sim\int_2^x\frac{\mathrm{d}t}{\log t}$.虽然右边和 $\frac{x}{\log x}$ 是同阶的函数 $\left(\int_2^x\frac{\mathrm{d}t}{\log t}\sim\frac{x}{\log x}\right)$,但是在 4×10^{16} 以内各种区间段的数值结果表明用 $\int_2^x\frac{\mathrm{d}t}{\log t}$ 作为 $\pi(x)$ 的近似值均比 $\frac{x}{\log x}$ 要好.一直到1896年解析数论产生并且有长足的进步之后,$\pi(x)\sim\frac{x}{\log x}$ 才由两位法国数学家Hadamard和de la Vallee Poussin各自独立地证明.

经典代数数论中的例子:高斯对二次域的理想类数作了大量数值计算,提出两个著名的猜想:

(1) 类数为1的虚二次域只有他算出的9个($\boldsymbol{Q}(\sqrt{-d}),d=$ 3,4,7,8,11,19,43,67,163).

(2) 类数为1的实二次域有无穷多个.第一个猜想已于1966年~1967年由Baker和Stark各自独立地证明,而第二个猜想至今未解决.

近代代数数论中的例子:1960年前后英国人Birch和Swinnerton-Dyer对椭圆曲线 E 的有理点群 $E(\boldsymbol{Q})$ 和 L-函数 $L_E(s)$ 作了大量的计算,提出关于 $E(\boldsymbol{Q})$ 的代数结构和 $L_E(s)$ 解析特性之间联系的重要猜想(见§3.6).近十年来这个猜想有很大进展,但未最后解决.

2.从统计学和其他理性思考提出猜想,然后用计算数据来考查猜想的正确程度.或者计算与推理相结合来证明猜想.

在解析数论中这方面有大量的例子.例如关于孪生素数的猜想是说:存在无限多个素数对($p,q=p+2$).如果以 $\pi_2(x)$ 表示 p

$\leqslant x$ 的孪生素数对$(p,q=p+2)$个数. 从统计学考虑,1923 年 Hardy 和 Littlewood 给出猜想

$$\pi_2(x)\sim C_2\frac{x}{(\log x)^2}\qquad (x\to+\infty),$$

其中 $C_2=\prod_{p\geqslant 3}\left(1-\frac{1}{(p-1)^2}\right)=0.66016\cdots$. 1976 年 Brent 计算出 $\pi_2(10^l)(l=3,4,\cdots,11)$,所得数据与上述猜想符合得很好,但是至今还不知道孪生素数对是否有无穷多对.

同样地,Hardy 和 Littlewood 在同一文章中还提出如下的猜想:

哥德巴赫问题:哥德巴赫于 1742 年写给欧拉的信中表示:每个整数 $n\geqslant 6$ 均可表示成三个素数之和,每个偶数 $2n\geqslant 4$ 均可表成两个素数之和. 若以 $r_k(n)$表示 n 表成 k 个素数之和的方法数, Hardy-Littlewood 猜想:当 $n\to+\infty$时,

$$r_2(2n)\sim C_2\frac{2n}{(\log 2n)^2}\prod_{3\leqslant p|n}\frac{p-1}{p-2},$$

其中 c_2 如上所示

$$r_3(n)\sim\frac{n^2}{2(\log n)^3}\prod_{p}\left(1+\frac{1}{(p-1)^2}\right)\prod_{3\leqslant p|n}\left(1-\frac{1}{p^2-3p+3}\right).$$

1964 年 Shen 用筛法计算到 $2n\leqslant 33\times 10^6$,证明偶数 $2n$ 均为两素数和. 1965 年 Stein 对 $2n\leqslant 2\times 10^5$ 算出 $r_2(2n)$的值,与上述渐近公式相当一致. 但是关于偶数表成二素数和的哥德巴赫问题至今未完全解决. 另一方面,维诺格拉多夫于 1938 年证明了:每个充分大的整数 n 均是三素数之和.

代数数论方面一个著名的例子是 Cohen 和 Lenstra 于 1984 年对于阿贝尔数域的类群为特定类型的有限群时,其阿贝尔数域所占的比例提出了猜想. 以虚二次域为例. 设 S 为所有奇阶有限阿贝尔群(的同构类)所成的集合,Aut G 表示 G 的自同构群. 对于

集合 S 上的每个非负实函数 f，令

$$M(f)=\lim_{n\to\infty}\sum_{|d_K|\leqslant n}f(c'(K))\Big/\sum_{|d_k|\leqslant n}1,$$

$$M_0(f)=\lim_{n\to\infty}\sum_{|G|\leqslant n}(f(G)/\mathrm{Aut}\ G)\Big/\sum_{|G|\leqslant n}(\mathrm{Aut}\ G)^{-1},$$

其中 $C'(k)$表示虚二次域 K 的理想类群的奇部分，$|d_k|\leqslant n$ 表示过判别式绝对值$\leqslant n$ 的所有虚二次域，$|G|\leqslant n$ 表示过阶数$\leqslant n$ 的所有奇阶阿贝尔群(同构类). Cohen-Lenstra 猜想是说：$M(f)=M_0(f)$.

取 T 为特定的奇阶有限阿贝尔群类型，定义 f 为 T 的特征函数，即 $f(G)=1$(若 $G\in T$)，$f(G)=0$(若 $G\notin T$). 则 $M_0(f)$可以具体算出来，而 $M(f)$是类群奇部分属于类型 T 的虚二次域在所有虚二次域中所占的比例. 取 T 为"所有奇阶循环群"，f 为 T 的特征函数，则可算出

$$M_0(f)=\frac{\zeta(2)\zeta(3)}{3\cdot(2)_\infty A\zeta(6),}\sim 0.977575,$$

其中 $A=\prod_{k\geqslant 2}\zeta(k)\sim 2.29486$，$\zeta(s)$为黎曼 zeta 函数，$(p)_r=\prod_{1\leqslant k\leqslant r}(1-p^{-k})$. 所以 Cohen-lenstra 猜想是说：类群奇部分是循环群的虚二次域所占的比例为 97.7575%. 类似方法可知这个猜想还说：对每个奇素数 p，类数被 p 除尽的虚二次域所占比例为 $f(p)=1-(p)_\infty=1-\prod_{k=1}^{\infty}(1-p^{-k})=\frac{1}{p}+\frac{1}{p^2}-\frac{1}{p^5}-\cdots$. 比如 $f(3)\approx 0.43987$，$f(5)\approx 0.23967$，$f(7)\approx 0.16320$.

类群的 p-部分同构于给定的有限阿贝尔 p-群 G 的虚二次域比例为$(p)_\infty/\mathrm{Aut}(G)$.

类群的 p-部分的秩为 r 的虚二次域所占比例为 $p^{-r^2}(p)_\infty/((p)_r)^2$.

对于其他类型的阿贝尔数域，Cohen 和 Lenstra 也给出类似的猜想，比如对实二次数域情形，他们猜想类数为 1 的实二次域所

占的比例为

$$\lim_{n\to\infty}\frac{\#\{\text{实二次域 } K\,|\,d_K\leqslant n, h(K)=1\}}{\#\{\text{实二次域 } K\,|\,d_K\leqslant n\}}\sim 0.75446,$$

即他们猜想大约有 3/4 的实二次域类数为 1. 这比高斯猜想(类数为 1 的实二次域有无穷多个)要强许多. 更一般地,对每个奇素数 p,他们猜想

p 除尽类数的实二次域所占比例为

$$1-\frac{(p)_\infty}{1-\frac{1}{p}}=\frac{1}{p^2}+\frac{1}{p^3}+\frac{1}{p^4}+\cdots$$

类群的奇数部分同构于给定奇阶阿贝尔群 G 的实二次域比例为 $m(G)=1/2gA(2)_\infty|\mathrm{Aut}G|$,其中 $g=|G|$. 例如 $m(\{1\})\approx 0.75446$, $m(\mathbf{Z}/3\mathbf{Z})\approx 0.12574$, $m(\mathbf{Z}/5\mathbf{Z})\approx 0.03772$.

类群的 p-部分的秩为 r 的实二次域比例为 $p^{-r(r+1)}(p)_\infty/(p)_r(p)_{r+1}$.

1987 年～1994 年,Cohen 和 J. Martinet 把猜想又推广到任意类型的数域上,并且表明这些猜想与实验数据惊人地一致. 但是要证明这些猜想现在看来是非常困难的.

计算与理论相结合,常常可以证明重要结果. 例如关于虚二次域的高斯猜想,早在 1934 年 Heilbronn 就证明出除了高斯算出的 9 个类数为 1 的虚二次域之外,至多还有一个这样的虚二次域,而且这个域若存在,其判别式$\leqslant 10^7$. 当时的计算技术和方法还不能把判别式$\leqslant 10^7$ 的所有虚二次域类数算出来. 后来于 1967 年 Baker 和 Sturk 各自采用超越数论和模形式理论,把上界 10^7 改小,才证明了例外虚二次域不存在. 类似地,近年来对一些类型的虚阿贝尔域(如虚四次循环域和伽罗华群同构于$(\mathbf{Z}/2\mathbf{Z})^2$ 的虚阿贝尔域等)决定出小类数(为 1 或 2)的全部数域,也是先用解析数论和代数方法对类数整除性和下界给出估计,然后对这样限制下来的有限个可能性进行逐个计算,最后得出结果.

另一个例子是关于 $\zeta(s)$ 的黎曼猜想. 1986 年,Van de Lune,

Riele 和 Winter 计算了 $\zeta(s)$的前 1.5×10^9 个(按绝对值)非平凡零点，它们均在直线 $\mathrm{Re}(s)=\frac{1}{2}$之上.

3. 应用领域的实际需要提出代数数论新的计算课题和算法改进.

代数数论的理论研究中提出了一系列传统的计算课题，比如计算各种类型数域的整基、类群结构和类数、基本单位系等，计算椭圆曲线有理点群和寻求秩很大的椭圆曲线，计算各种模形式的 Fourier 系数和 Hecke 算子在模形式空间中的本征值等. 这方面的算法在不断改进，算出的数据表也在不断地扩大. 新的算法和数据已制成软件供人们使用. 另一方面，20 世纪后半期计算机科学和信息工程等方面的实际需要提出了大量新的计算课题，极大地促进了计算代数数论的发展. 例如：

(1)20 世纪 50 年代末期计算机和通信系统采取移位寄存器，容易实现有限域上四则运算和递归运算，有限域成为重要的代数工具，并且从实用角度提出了许多有限域上的计算问题，如：寻求有限域上多项式因子分解的快速算法，构作有限域上(首 1)不可约多项式和本原多项式表($\boldsymbol{F}_q$ 上 n 次首 1 多项式叫作本原的，是指它的根为 $\boldsymbol{F}_{q^n}$中 q^n-1 阶元素)，寻求 $\boldsymbol{F}_{q^n}$对 $\boldsymbol{F}_q$ 的正规基$\{\alpha,\alpha^q,\alpha^{q^2},\cdots,\alpha^{q^{n-1}}\}$，正交基$\{\alpha_1,\cdots,\alpha_n\}$(即当 $i\neq j$ 时 $T_r(\alpha_i\alpha_j)=0$)和自对偶基$\{\alpha_1,\cdots,\alpha_n\}$(即 $T_r(\alpha_i\alpha_j)=\delta_{ij}$)，寻求 $\boldsymbol{F}_q$ 上次数最小的多项式 $f(x)$，$f(0)=1$，使得有 $g(x)\in\boldsymbol{F}_q[x]$，$\deg g(x)<\deg f(x)$，并且幂级数 $\frac{g(x)}{f(x)}=\sum\limits_{i=0}^{\infty}c_ix^i\in\boldsymbol{F}_q[[x]]$ 的前 n 个系数 $c_i(0\leqslant i\leqslant n-1)$为给定值(这叫线性移位寄存器的综合算法. 最好算法是 Massey 于 1965 年得到的，用于 BCH 码的纠错译码)，构作有限域上具有良好相关性能的周期序列和周期序列组(用于多址通信的同步)，构作从 $\boldsymbol{F}_q^n$ 到 $\boldsymbol{F}_q^m$ 的函数使其满足某些特定的条件(Hash 函数，Bent 函数，相关免疫函数等). 对于后两类问题，代数数论(特

别是函数域和高维代数簇的算术)已成为重要工具.关于代数曲线的韦依定理和高维代数簇的德林定理,自守表示 L-函数的零点特性被用来计算和估计各种指数和,而这些指数和的估计和计算被用来构作有限域上理想的周期序列组和有限域上理想的函数.

(2) 素性判定和大数分解问题.

在这个问题上充分体现出纯粹数学家和应用数学家对于数学具有不同的美学标准.公元前3世纪欧几里德证明了每个整数≥2惟一地表示成有限个素数之乘积.并且证明了素数有无限多个.纯粹数学家认为这些结论(包括证明)都是漂亮的,但是应用数学家却问:我需要一个100位的素数(注记:切夫雪夫于1850年证明了,对每个实数 $x>1$,在 x 和 $2x$ 之间必有素数),或者要把一个200位的整数具体进行素因子分解,是否有好的办法?

初等数论中有不少判定素数的办法.例如:正整数 $N\geqslant 2$ 是素数当且仅当 $(N-1)!\equiv -1 \pmod N$,这叫威尔逊定理.当 N 很大时,用这种办法判别 N 是否为素数计算量太大.又如我们有费马小定理:若 N 为素数,则对每个与 N 互素的整数 b,$b^{N-1}\equiv 1 \pmod N$.但这只是素数的一个必要条件.不是充分条件.最小的反例为 $561=3\cdot 11\cdot 17$.进一步则用欧拉引理:若 n 为奇整数,如果 n 是素数,则对与 n 互素的整数 b,

$$b^{\frac{n-1}{2}}\equiv\left(\frac{b}{n}\right) \pmod n, \tag{1}$$

其中 $\left(\frac{b}{n}\right)$ 是雅柯比符号,即若 $n=p_1\cdots p_l$ 为 l 个素数之积,则 $\left(\frac{b}{n}\right)=\left(\frac{b}{p_1}\right)\cdots\left(\frac{b}{p_l}\right)$,其中 $\left(\frac{b}{p_i}\right)$ 是朗让得符号.雅柯比符号 $\left(\frac{b}{n}\right)$ 不必用 n 的素因子分解式,可以用与勒让得符号类似的互反律来运算,计算量为 $O((\log n)^3)$,$b^{\frac{n-1}{2}} \pmod n$ 的计算量也为 $O((\log n)^3)$.另一方面可以证明:若 n 不是素数,则至少有一半的 $b\in(\mathbf{Z}/nZ)^x$,使得(1)式不成立.于是对任意大的奇整数 n,我们随机地选取 $b\in(\mathbf{Z}/n\mathbf{Z})^x$(为试验是否 $(b,n)=1$,可用辗转相除法计算 (b,n),计算

量为 $O(\log n)$)，看试验(1)是否能通过(即公式(1)是否成立). 如果对某个 b 不能通过试验(1)，则 n 为合成数. 如果试验(1)对 m 个 b 均能通过，则 n 是合成数的概率不超过$\left(\frac{1}{2}\right)^m$，从而 n 为素数的概率大于 $1-\left(\frac{1}{2}\right)^m$. 这样的算法叫作概率算法，是 1977 年 Solovay 和 Strassen 给出的.

再进一步的素数判定算法是 1980 年的 Miller-Rabin 试验方法：对正奇数 n，令 $n-1=2^s m$(m 为奇数，$s\geqslant 1$). 如果 n 为素数，易知对每个 $b\in(\mathbf{Z}/n\mathbf{Z})^x$，下面论断成立

或者 $b^m\equiv 1 \pmod n$，

或者有 $1\leqslant t<s$，使 $b^{m\cdot 2^t}\equiv -1 \pmod n$. (2)

可以证明(用解析数论)：

(A) 若 n 不为素数，则最多有 25% 的 $b\in(\mathbf{Z}/n\mathbf{Z})^x$ 通过试验(2).

(B) 如果对于狄里赫利 L-函数 $L(s,\chi)$的广义黎曼猜想成立(即 $L(s,\chi)$的非平凡零点都在直线 $\mathrm{Re}(s)=\frac{1}{2}$之上)，则 n 为素数当且仅当对所有 $b\leqslant 2(\log n)^2$ 都能通过试验(2).

于是我们对每个 $b\in(\mathbf{Z}/n\mathbf{Z})^x$ 试验条件(2). 若对某个 b 不能通过试验(2)，则 n 为合成数. 若对 m 个 b 均通过试验(2)，则 n 至少以 $1-\left(\frac{1}{4}\right)^m$ 的概率为素数，这为概率算法. 如果广义黎曼猜想成立，则当 b 从 2 一直试验到$[2(\log n)^2]$时均能通过试验(2)，则 n 为素数，这是"决定"型的算法.

此后，素数判定算法仍不断改进，它们都是多项式算法，即判别 n 是否为素数的计算量为 $\log n$ 的多项式，例如利用高斯和的 Adleman-Pomerance-Rumely 方法(1983 年)和 Cohen-Lenstra 方法(1984 年). 对于某些特殊类型的整数，可以给出更好的素数判定算法. 例如对费马数 $F_n=2^{2^n}+1$，欧拉就证明了 F_n 的因子必有

形式 $k \cdot 2^{n+2}+1$. 他对 F_5 试验 $k=5$，发现了 F_5 的一个素因子 $5 \cdot 2^7+1=641$. 又如对 Mersenne 数 $M_q=2^q-1$(其中 q 为素数)，$M_2=3$，$M_3=7$，$M_5=31$，$M_7=127$ 均为素数，但是 $M_{11}=23\times 89$. 欧拉叙述(1750 年)而由拉格朗日(1775 年)证明了以下结果：若素数 $q\equiv 3(\bmod 4)$，则 $2q+1 \mid M_q$ 当且仅当 $2q+1$ 为素数. 由此得出许多 Mersenne 合成数. 1952 年 Robinson 和 Lehmer 用计算机发现 M_{521}，M_{607} 为素数，这是历史上第一批用计算机发现的素数.

寻求大素数的问题是由于信息技术的需要(见 § 5.3)，用目前最好的计算机(硬件)和最好的素数判定算法，判定一个 200 位左右的数是否为素数只需几分钟. 另一方面，将一个大整数完全分解成素因子乘积则要困难得多(注意：前面所有的算法在判定出 n 为合成数时，并没有得出 n 的真因子). 目前对一个 200 位左右的整数作素因子分解，估计需要万年以上. 目前一种普遍应用的密码体制就是建立在大数分解非常困难的信念上(见 § 5.3). 但这种信念是否正确，就导致近 20 年来人们对研究大数分解新算法的强烈兴趣. 1984 年，美国计算机协会(ACM)为纪念美国电子工程学会(IEEE)一百周年作了一个纪念牌，上面刻有用二次筛法完成的 $2^{251}-1$ 的素因子分解式. ACM 主席还作了以下的注解：

“大约 300 年前法国数学家 Mersenne 预言 $2^{251}-1$ 是合成数. 大约 100 年前证明了它是合成数，但直到 20 年前还被认为没有进行分解的计算装置. 事实上，用通常计算机和传统算法，估计其计算时间是 10^{20}年. 今年 2 月这个数在 Sandia 的 Cray 计算机上用 32 小时被分解成功，这是一个世界记录. 我们在计算方面已走了很长的路程. 为了纪念 IEEE 对计算的贡献，在这里刻上这个 Mersenne 数的 5 个素因子.”

300 多年前，费马就用平方差方法求整数的因子：如果 n 可表成平方差 a^2-b^2，则有分解 $n=(a+b)(a-b)$. 换个角度说：每个奇的合成数 $n=ab$ 均可表成平方差 $n=\left(\frac{a+b}{2}\right)^2-\left(\frac{a-b}{2}\right)^2$. 我们可

以依次取 $A=[\sqrt{n}]+1,[\sqrt{n}]+2,\cdots$，看 A^2-n 是否为完全平方数. 但这种方法只有 n 的两个因子 a 和 b 相差不多时才比较有效，这时 $\frac{a+b}{2}$ 接近于 n. 到了 20 世纪 20 年代，Kraitchik 改进了费马的平方差技术，这种改进成为目前最现代的分解算法的基础.

Kraitchik 的想法是：不必去寻找 u 和 v 使得 u^2-v^2 恰好为 n，而是认识到只要使 u^2-v^2 为 n 的倍数（即 $u^2\equiv v^2 \pmod{n}$）就可能成功. 如果奇数 n 为合成数，则 $u^2\equiv v^2 \pmod{n}$ 至少有 $\frac{1}{2}$ 的可能性使得 $u\not\equiv\pm v \pmod{n}$. 这时 $(u-v,n)$ 一定是 n 的非平凡因子.

以 $n=2041$ 为例，$[\sqrt{n}]+1=46$. 对 $Q(x)=x^2-2041$ 分别取 $x=46,47,\cdots$，得到 $Q(x)=75,168,263,360,459,560,\cdots$ 它们全不是平方数. 所以若用费马方法还要继续下去. 而 Kraitchik 的想法是寻求一批 x_i，使乘积 $Q(x_1)\cdots Q(x_k)$ 为平方数 v^2，令 $x_1\cdots x_k=u$，则

$$u^2=x_1^2\cdots x_k^2\equiv Q(x_1)\cdots Q(x_k)=v^2 \pmod{n},$$

就得到整数 u,v，使 $u^2\equiv v^2 \pmod{n}$.

问题是：如何求适当的 $x_1,\cdots,x_k$? Kraitchik 发现有些 $Q(x)$ 的分解是容易的：对于 $x=46,47,49$ 和 51，$Q(x)=x^2-2041$ 分别可分解为

$$75=3\times 5^2,168=2^3\times 3\times 7,360=2^3\times 3^2\times 5,560=2^4\times 5\times 7,$$

它们相乘是平方数. 于是 $u^2\equiv v^2 \pmod{2041}$，其中

$$u=46\cdot 47\cdot 49\cdot 51\equiv 311 \pmod{2041},$$

$$v=2^5\cdot 3^2\cdot 5^2\cdot 7\equiv 1416 \pmod{2041}.$$

由于 $311\not\equiv\pm 1416 \pmod{2041}$，算出 $(1416-311,2041)=13$，得到分解 $2041=13.157$.

基于上述思想，Brillhart 和 Morrison 于 1975 年给出如下的算法. 取定一个适宜的正整数 B，设 $p_1,p_2,\cdots,p_B$ 是前 B 个素数（叫作“因子基”）. 我们试图寻找一些整数 x_i，使得 $Q(x_i)=x_i^2-n$

的素因子均不超过 p_B. 于是 $Q(x_i)=p_1^{a_1}p_2^{a_2}\cdots p_B^{a_B}$，然后把它对应一个二元域上长为 B 的向量 $V(Q(x_i))=(\bar{a}_1,\bar{a}_2,\cdots,\bar{a}_B)\in \boldsymbol{F}_2^B$. 如果找到充分多这种 x_i（比如找到 $B+1$ 个），则对应的向量在 $\boldsymbol{F}_2$ 上线性相关，于是某些向量 $V(Q(x_i))(1\leqslant i\leqslant t)$ 在 $\boldsymbol{F}_2$ 上之和是全零向量，从而 $Q(x_1)\cdots Q(x_t)$就是完全平方数. 以上面的 $n=2041$ 为例，取 $B=4$，则因子基为 2,3,5,7. 我们有

$$V(75)=V(Q(46))\equiv(0100)\pmod 2,$$
$$V(168)=V(Q(47))\equiv(1101)\pmod 2,$$
$$V(360)=V(Q(49))\equiv(1010)\pmod 2,$$
$$V(560)=V(Q(51))\equiv(0011)\pmod 2,$$

这 4 个向量之和为零向量，从而 75 • 168 • 360 • 560 为平方数.

B 的选取要适当，B 选取太小则寻找 x_i 使 $Q(x_i)$的所有素因子均不超过 p_B 太困难，而 B 太大则需要得到很多 $Q(x_i)$才能使对应向量 $V(Q(x_i))$之和为零. 通过解析数论的估算，当取 B 使 p_B 大约为 $\exp\left(\sqrt{\frac{1}{2}\log_n\log\log n}\right)$ 的时候，用上述方法得到 n 的一个真因子的计算量达到最小值 $\exp(\sqrt{2\log n\log\log n})$. 所以这是一个比指数算法要好的“亚指数”算法，但还不是（对 $\log n$ 的）多项式算法.

差不多同时代还有采用同样思想的连分数算法. 后来又出现了二次筛法算法（Pomerance，1982），类群方法（Schnorr 和 Lenstra，1984），椭圆曲线算法（Lenstra，1987）和数域筛法（Lenstra，1993），后三种算法利用了数域中类群结构和素理想分解以及有限域上椭圆曲线的点群结构，是目前最有效的大数分解算法，但计算量也均是亚指数型的.

(3) 离散对数问题.

对每个有限域 $\boldsymbol{F}_q$，乘法群 $\boldsymbol{F}_q^\times$ 是 $q-1$ 阶循环群. 设 g 是它的一个生成元，则 $\boldsymbol{F}_q^\times$ 中每个元素 a 都可表成 g^e，$0\leqslant e\leqslant q-2$. 由 g 和指数 e 计算 a 是容易的，计算量为 $\log e$ 的多项式函数（多项式算

法)，但是由 a,g 计算 e 被认为是困难的，这叫作**离散对数**问题. 更一般的，对于任意有限阿贝尔群 G 和 G 中两个元素 a 和 b，是否有整数 e 使得 $a=b^e$? 如果是的话如何求 e? 对许多具体群 G 目前只有亚指数型的算法. 比如对于数域的理想类群，函数域的除子类群，有限域上椭圆曲线的点群等情况都是这样. 基于这种状况，这些群上的离散对数问题的困难性也被用于保密通信的许多方面. 由于需要相应的群结构比较复杂，所以也同时引发出一系列研究课题，如对于有限域 $\boldsymbol{F}_q$ 如何构作其上的椭圆曲线，使得点群有阶数很大的元素. 这样的构作性问题往往需要椭圆曲线理论中相当深刻的结果.

以上我们对计算代数数论的发展和现状作一简单的介绍. 其研究课题和风格与纯粹代数数论研究有所不同，但是现代代数数论的许多理论成果在应用和计算领域发挥了很大的威力.

参 考 文 献

[1] E. Bach and J. Shallit, Algorithmic Number Theory, vol. 1: Efficient Algorithms, MIT Press, 1996.

[2] H. W. Lenstra and R. Tijdeman (eds.), Computational Methods in Number Theory, Math. Centre tracts 154/155, Math. Centrum Amsterdam, 1982.

[3] A. Petho, M. Pohst, H. Williams and H. G. Zimmer (eds.) Computational Number Theory, Walter de Gruyter, 1991.

[4] M. Pohst (ed.), Algorithmic Methods in Algebra and Number Theory, Academic Press, 1987.

[5] M. Pohst and H. Zassenhaus, Algorithmic Algebraic Num-

ber Theory,Cambridge Univ. Press,1989.

[6] M. Pohst,Computational Algebraic Number Theory,DMV Seminar 21,Birkhauser,1993.

[7] H. Zimmer,Computational Problems,Methods and Results in Algebraic Number Theory,LN in Math. 262,Springer-Verlag,1972.

[8] H. Cohen, A Course in Computational Algebraic Number Theory,GTM 138,Springer-Verlag,1996.

[9] E. Berlekamp, Factoring polynomials over large finite fields,Math. Comp. 24(1970),713～735.

[10] D. Cantor and H. Zassenhaus,A new algorithm for factoring polynomials over finite fields,Math. Comp. 36(1981), 587～592.

[11] H. Cohn and H. W. Lenstra,Heuristics on class groups of number fields,LN in Math. 1068,Springer-Verlag,1984.

[12] H. Cohn and J. Martinet,Class groups of number fields:numerical heuristics,Math. Comp. 48(1987),123～137.

[13] —and—,Etude heuristiquedes groupes de classes des corps de nombres,J. Reine Angew. Math. 404(1990),39～76.

[14] —and—,Heuristics on class groups:some good primes are not too good,Math. Comp. 63(1994),329～334.

[15] S. Fermigier,Un example de courbe elliptique definie sur $\boldsymbol{Q}$ de rang≥19,C. R. Acad. Sci. Paris,315(1992),719～722.

[16] K. Nagao and T. Kouya,An example of elliptic curve over $\boldsymbol{Q}$ with rank≥21,Proc. Japan Acad. ,70(1994),104～105.

[17] D. Wiedemann,Solving sparse linear equations over finite fields,IEEE Trans. Information Theory 32(1986),54～62.

§5.2 代数几何码

在现代数字通信技术时代，各种具体信息（电报，电话，图像，数据…）都要编成数字符号传输，每个信息编成某有限集合 S 中元素的一个元素组.最常见的是 $S=\{0,1\}$（二元域），也常把 S 取成有限域 $\boldsymbol{F}_q$，因为 $\boldsymbol{F}_q$ 中可以作四则运算.

假如我们有八个信息，在二元域 $\boldsymbol{F}_2$ 上被分别编成(000)，(001)，(010)，(011)，(100，(101)，(110)，(111)传输.如果(011)在信道传输中产生错误，比如第 2 位的 1 错成 0，那么收方得到(001)，并且毫不知晓这是错的（检错），更无法知道发方传给他的是(011)（纠错），因为收到的(001)也是有意义的.

为了使通信有检错和纠错能力，一个最简单的办法是把每个码(abc)重复三次编成($abcabcabc$)再传出.所以有意义的信号组成 $\boldsymbol{F}_2$ 上 q 维向量空间（2^9 个向量）中一个 8 元子集合

$$C=\{(abcabcabc)\mid a,b,c\in\boldsymbol{F}_2\}.$$

发出信息(001001001)之后，若在信道中出两个错，收到的都不是 C 中元素，所以可以检两个错.若在传输中出一个错，例如第 1 位出错，收到向量(101001001)，我们可以把错纠正过来，因为在集合 C 中只有原始信息与它相差一位，而 C 中其它元素与(101001001)至少相差两位，从而把(101001001)译成和他相差位最少的 C 中惟一元素(001001001).于是可以纠一个错.但是另一方面，本来用三位就可传输 1 个信息，现在要用九位，效率（传输速度）变成原来的 1/3.所以是靠牺牲了传输效率来提高纠错能力的.

一般的，$\boldsymbol{F}_q$ 上 n 维向量空间 $V=\boldsymbol{F}_q^n$ 的任何子集合 C 都叫一个

码，C 中元素叫作码字，n 叫码长，C 中元素个数表成 $K=|C|$，这 K 个信息本来用 $k=\log_q K$ 位就可以传输，k 叫信息位数，但是现在用了 n 位，所以传输效率为 $\rho=\frac{k}{n}$.

为了衡量码 C 的纠错能力，我们对 V 中每个向量 $x=(x_1,\cdots,x_n)(x_i\in \boldsymbol{F}_q)$ 定义 Hamming 权为

$$W_H(x)=\{i\mid 1\leqslant i\leqslant n, x_i\neq 0\}(\text{非零分量的个数}).$$

而对任意两个向量 x,y，定义它们的 Hamming 距离为 $d_H(x,y)=W_H(x-y)$（即 x 和 y 的相异位数）. d_H 确实是通常严格数学意义下的距离，即它满足以下三条件：对 $x,y,z\in V$，

(1) $d_H(x,y)\geqslant 0$，并且 $d_H(x,y)=0$ 当且仅当 $x=y$；

(2) （对称）$d_H(x,y)=d_H(y,x)$；

(3) （三角形不等式）$d_H(x,z)\leqslant d_H(x,y)+d_H(y,z)$.

最后，定义码 C 的最小距离为

$$d=d(C)=\min\{d_H(x,y)\mid x,y\in C, x\neq y\}.$$

纠错码理论建立在下面一个简单事实的基础之上.

基本定理 设 $d(C)=d$，$e=\left[\frac{d-1}{2}\right]$，则码 C 可检查出 $\leqslant d-1$ 个错，可纠正 $\leqslant e$ 个错.

证明 设 $x\in C$，ε 为错误向量，接收到 $y=x+\varepsilon$. 则 $d_H(x,y)=W_H(y-x)=W_H(\varepsilon)$. 若 $1\leqslant W_H(\varepsilon)\leqslant d-1$，则 y 不为码字，从而可检查出 $\leqslant d-1$ 个错. 若 $1\leqslant W_H(\varepsilon)\leqslant e$，由三角形不等式知对每个码字 $x'\in C$，$x\neq x'$，则 $d\leqslant d_H(x,x')\leqslant d_H(x,y)+d_H(x',y)\leqslant \varepsilon+d_H(x',y)$，于是 $d_H(x',y)\geqslant d-e\geqslant e+1>d_H(x,y)$. 这表明 x 是惟一的码字与接收向量 y 距离达到最小值. 把 y 译成码字 x，从而可纠 $\leqslant e$ 个错.

于是，从纠错码的角度，$\boldsymbol{F}_q$ 上的一个码 C 有三个重要参数：码

长 n，信息位数 $k=\log_q|C|$ 和最小距离 $d=d(C)$. 一个好的纠错码 C 是希望有小的 n 值和大的 k,d 值，但这三者之间是相互制约的，例如对固定的码长 n，如果 k 很大(即码 C 中码字很多)，就不能期望有太大的最小距离 d. 所以纠错码理论的基本问题为：

(1) 给了三个参数 n,k,d 当中的两个，如何估计第三个参数的界？如何构作出第三个参数最优的码 C？

第一个问题是估计纠错码的纠错能力和传输效率能够好到什么程度，而第二个问题则是工程师们更为关心的，即如何具体给出好的纠错码供通信系统使用. 此外，工程师们也关心这些好的纠错码在技术上是否容易实现把接收向量 y 译成发出码字 x 的纠错译码. 让我们看一下上述基本定理的证明：为了从 y 恢复成 x，需要把 C 中所有码字均与 y 比较，从中找出与 y 距离最小的惟一码字 x. 这个算法显然是太耗时了. 所以从工程角度考虑，纠错码理论的另一个重要课题是

(2) 对于构作出的好的纠错码 C，寻求可技术实现的好的译码算法.

40 年来，为了使用更好的数学工具来研究纠错码，人们不仅从集合论或组合论的观点研究码的性质，还需赋以代数结构. 首先人们考虑 C 是 $V=\boldsymbol{F}_q^n$ 在 $\boldsymbol{F}_q$ 上的向量子空间，这叫作**线性码**，这时信息位数 k 就是子空间 C 的维数. 而最小距离 $d(C)$ 就是 C 中非零码字的最小 Hamming 权. 纠错编码和译码都可用矩阵工具. 若线性码 C 还满足条件：如果 $(a_1,\cdots,a_n)\in C$，则 $(a_n,a_1,a_2,\cdots,a_{n-1})\in C$. 则 C 叫作**循环码**. 这时，若把码字 $c=(c_0,c_1,\cdots,c_{n-1})\in C$ 对应于有限环 $R=\boldsymbol{F}_q[x]/(x^n-1)$ 中的元素 $c_0+c_1x+\cdots+c_{n-1}x^{n-1}$，则整个码 C 就是环 R 中的一个理想. 而 R 是主理想环，所以每个理想有形式 $(f(x))=f(x)R$，其中 $f(x)$ 是 x^n-1 在 $\boldsymbol{F}_q[x]$ 中的因子，于是对循环码的研究可以使用近世代数工具.

比如设 C 是 $\boldsymbol{F}_q$ 上的线性码，即 C 是 $V=\boldsymbol{F}_q^n$ 的 k 维子空间，以 $v_i=(a_{i1},\cdots,a_{in})(1\leqslant i\leqslant k)$ 表示 C 的一组基，则 $\boldsymbol{F}_q$ 上的 $k\times n$ 矩阵

（秩为 k）

$$G=\begin{pmatrix}V_1\\ \vdots\\ V_k\end{pmatrix}=\begin{pmatrix}a_{11}\cdots a_{1n}\\ \cdots\\ a_{k1}\cdots a_{kn}\end{pmatrix}$$

叫作码 G 的生成矩阵. 于是 $C=\{xG\mid x=(x_1,\cdots,x_k)\in \boldsymbol{F}_q^k\}$，即把 $\boldsymbol{F}_q^k$ 中 q^k 个信息 x 编成码长为 n 的码字 xG，这就是纠错编码. 另一方面，在 V 中引入内积：

$$\langle x,y\rangle=x_1y_1+\cdots+x_ny_n\in \boldsymbol{F}_q, x=(x_1,\cdots,x_n),$$
$$y=(y_1,\cdots,y_n),$$

则 $C^{\perp}=\{y\in V\mid 对每个\ x\in C, \langle x,y\rangle=0\}$ 也是 V 中一个线性码，维数为 $n-k$，叫作 C 的对偶码. 以 $u_j=(b_{j1},\cdots,b_{jn})(1\leqslant j\leqslant n-k)$ 表示 $C^{\perp}$ 的一组基，则 $(n-k)\times n$ 矩阵（秩为 $n-k$）

$$H=\begin{pmatrix}u_1\\ \vdots\\ u_{n-k}\end{pmatrix}=\begin{pmatrix}b_{11} & \cdots & b_{1n}\\ & \cdots & \\ b_{n-k,1} & \cdots & b_{n-k,n}\end{pmatrix}$$

叫作 C 的校验矩阵. 这是由于对每个 $y\in V$，$y\in C$ 当且仅当 $yH^T=0$（长为 $n-k$ 的零向量）. 由此不难看出，线性码 C 的最小距离 d 是最小正整数，使得 H 的任 $d-1$ 列都 $\boldsymbol{F}_q$-线性无关，但是存在 d 列是 $\boldsymbol{F}_q$-线性相关的. 所以利用这个性质可以由 H 决定线性码 C 的最小距离，还可由此设计校验矩阵 H 以得到 d 很大的线性纠错码 C. 又若码字 x 经传输产生错误 ε 而收到向量 $y=x+\varepsilon$，如果 $1\leqslant W_H(\varepsilon)\leqslant d-1$，则 $yH^T=(x+\varepsilon)H^T=\varepsilon H^T\neq 0$. 所以只需计算 yH^T 就可知道 y 不是码字，发现有错. 如果 $W_H(\varepsilon)=e, 2e+1\leqslant d$，则 $yH^T=\varepsilon H^T$ 惟一地表示成 H 中 e 列的线性组合. 设 yH^T 是 H 的第 $i_1,\cdots,i_e$ 列的线性组合，系数分别为 $c_{i_1},\cdots,c_{i_e}$，则错误向量 ε 在第 $i_1,\cdots,i_e$ 位分别为 $c_{i_1},\cdots,c_{i_e}$，而其余位为 0，所以通过计算 yH^T 可以纠错. 这就是利用矩阵进行线性码纠错的译码算法.

1960 年，Bose，Chaudhui 和 Hocquenghem 利用有限域的知识构作出一种线性码（并且常常基本上是循环码），后人称作 BCH

码. 这种码可以设计出具有任意大的最小距离. 1965 年, Berlekamp 和 Massey 各自独立地给出这种码非常好的译码算法, 用已有的线性移位寄存器技术很容易实现. 所以从那时起 BCH 码在通信系统中一直沿用至今, 尽管这种码的性能在理论上并不是最优的, 它不能达到 1952 年 Varshamov 和 Gilbert 给出的渐近界. 20 世纪 70 年代后期, 前苏联数学家 Goppa 从另一个角度审视 BCH 码, 然后于 1981 年用代数数论和代数几何把它作了极大的推广, 这就是下面的代数几何码.

设 C 是 $\boldsymbol{F}_q$ 上一条不可约代数曲线, 亏格为 $g=g(C)$, $K=\boldsymbol{F}_q(C)$ 为它的函数域. 对 K 的每个除子 D, 定义

$$L(D)=\{f\in K^x \mid \mathrm{div}(f)\cdot D \text{ 为整除子}\}\cup\{0\}.$$

如果 $D=\Pi p^{e_P}$, 其中 P 过 K 的所有素除子, 而除了有限个 P 之外, e_P 均为 0, 则 K 中非零函数 f 属于 $L(D)$ 当且仅当对 K 的每个素除子 P, $V_P(f)\geqslant -e_P$. 由 P 的非阿特性易知 $L(D)$ 是 $\boldsymbol{F}_q$ 上的向量空间. 由黎曼-洛赫定理可推出 $L(D)$ 是 $\boldsymbol{F}_q$ 上的有限维向量空间.

K 的每个 1 次素除子 P 对应于曲线 C 上的一个 $\boldsymbol{F}_q$-点, 这个点也记成 P. 如果 P 不是 K 中某函数 f 的极点(即 $V_P(f)\geqslant 0$), 则 f 在点 P 处的取值记为 $f(P)\in\boldsymbol{F}_q$.

现在设 $P_1,\cdots,P_n$ 是 K 的 n 个不同的 1 次素除子(它们对应曲线 C 上 n 个不同的 $\boldsymbol{F}_q$-点), $D=P_1\cdots P_n$. 又令 G 为 K 的一个整除子, 并且 G 和 D 没有公共素除子因子, 则 $L(G)$ 中元素 f 在 P_i 处均没有极点, 于是 $f(P_i)\in\boldsymbol{F}_q(1\leqslant i\leqslant n)$. 由黎曼-洛赫定理可证明如下结果.

定理(Goppa, 1981). 在上述假定下, 对于 $\boldsymbol{F}_q$ 上的线性码(码长为 n)

$$L_C(D,G)=\{(f(P_1),\cdots,f(P_n))\in\boldsymbol{F}_q^n \mid f\in L(G)\},$$

则当 $\deg G<n$ 时, 信息位数 $\geqslant \deg G-g+1$, 最小距离 $\geqslant n-\deg G$.

并且当 $2g-2<\deg G<n$ 时，信息位数 $=\deg G-g+1$.

以 $\Omega_C(D,G)$ 表示线性码 $L_C(D,G)$ 的对偶码，则在 $\deg G>2g-2$ 时，$\Omega_C(D,G)$ 的信息位数 $\geqslant n-\deg G+g-1$，最小距离 $\geqslant \deg G-2g+2$. 并且当 $2g-2<\deg G<n$ 时，$\Omega_C(D,G)$ 的信息位数 $=n-\deg G+g-1$.

注记：事实上，线性码 $\Omega_C(D,G)$ 可由函数域上的微分来给出更直接的定义.

$L_C(D,G)$ 和 $\Omega_C(D,G)$ 都叫作代数几何码. 它的构作方式非常灵活，我们可以选取各种代数曲线 C，C 上的 $\boldsymbol{F}_q$-点 $P_1,\cdots,P_n$ 和整除子 D. 1982 年，Tsfasman、Vladut 和 Zink 三人使用模曲线构作代数几何码，第一次得到了超过 Varshamov-Gilbert 渐近界的线性码. 这是纠错码理论的一个重大的突破，证明用到模形式和椭圆曲线理论近 30 年来的理论成果. 自那时起的十多年来，代数几何码的研究引起数论和代数几何学家的极大兴趣，纷纷对各种代数曲线（椭圆曲线，Hermite 曲线，费马曲线，超椭圆曲线等）研究代数几何码的结构（明显构作生成矩阵或校验矩阵，决定最小距离的值，决定码的自同构群以及何时为循环码等）. 1995 年，Garçia 和 Stichtenoth 用 Artin-Schreier 曲线这种更为初等的方法构作出超过 Varshamov-Gilbert 界的代数几何码. 在所有这些研究中，人们最为关注的是代数几何码的纠错译码算法. 目前译码算法已有很大进展. 一旦译码算法完全达到技术上实用的程度，它将取代 BCH 码，而通信工程师们将要学习代数数论和代数几何知识.

从 Goppa 定理看出，为了构作性能好的代数几何码，需要 n 很大，即曲线 C 在 $\boldsymbol{F}_q$ 上有很多点，因为对于码 $L_C(D,G)$，$d\geqslant n-\deg G$，而对于它的对偶码 $\Omega_C(D,G)$，传输效率为 $\frac{k}{n}\geqslant\frac{n-\deg G+g-1}{n}$. 于是从代数几何码的研制中产生了一个纯理论

问题：对于亏格为给定值 g 的所有 $\boldsymbol{F}_q$ 上曲线 C，点集合 $C(\boldsymbol{F}_q)$ 最大可能是多少？以 $N_q(C)$ 表示 $\boldsymbol{F}_q$ 上代数曲线 C 的 $\boldsymbol{F}_q$-点个数，$N_q(g)$ 表示亏格为 g 的曲线 C 的 $\boldsymbol{F}_q$-点最多个数. 即

$$N_q(g)=\max\{N_q(C)\mid g(C)=g\}.$$

再令 $A(q)=\overline{\lim\limits_{g\to\infty}}\dfrac{N_q(g)}{g}$. 问题是决定 $N_q(g)$ 和 $A(q)$ 的值. 这个问题引起了许多数学家(包括一些著名数学家 Serre，Ihara，Manin，Drinfeld 等人)的兴趣. 根据韦依定理，$|N_q(C)-(q+1)|\leqslant 2g\sqrt{q}$. Serre 又作了改进 $|N_q(C)-(q+1)|\leqslant 2g[\sqrt{q}]$，由此可知 $A(q)\leqslant 2[\sqrt{q}]$. 1983 年，Drinfeld 和 Vladut 把它又改进为 $A(q)\leqslant\sqrt{q}-1$. 而 1982 年 Tsfasman 等三人和 1995 年 Garcia 二人分别用模曲线和 Artin-Schreier 曲线构作代数几何码的工作，实际上证明了当 q 为平方数时 $A(q)\geqslant\sqrt{q}-1$. 所以对于 q 为平方数的情况已经决定出 $A(q)=\sqrt{q}-1$. 对于 q 为非平方数的情形这个问题还未解决，下界 $A(q)\geqslant c\log q$ 是由 Serre 得到的. Serre 下界虽然与上界相差甚远，但却用了很高深的数论方法：函数域的类域塔.

我们知道，数域 K 的最大不分歧阿贝尔扩张 K_1 叫作 K 的希尔伯特类域，而类域论表明 K_1/K 的伽罗华群同构于 K 的理想类群，所以 $[K_1:K]=h(K)$(K 的理想类数). 进而，以 K_2 表示 K_1 的希尔伯特类域，则 $[K_2:K_1]=h(K_1)$. 一般地，以 K_{n+1} 表示 K_n 的希尔伯特类域，则 $[K_{n+1}:K_n]=h(K_n)$. 希尔伯特在《数论报告》中问：是否对某些数域 K 存在无限的类域塔，即每个 K_{n+1} 都大于 K_n(这也相当于说 $h(K_n)$ 均 >1)？这个问题在 1964 年由前苏联数学家波列维奇和沙弗列维奇解决，他们对某些二次域 K 证明了类域塔是无限扩大的. Serre 把这种方法用到函数域上，得到 $A(q)$ 的下界. 以后有许多人用此方法得到了 $N_q(g)$ 的上下界优于韦依结果 $(q+1)-2g\sqrt{q}\leqslant N_q(g)\leqslant(q+1)-2g\sqrt{q}$，并且对某些 q 和 g

决定出 $N_q(g)$ 的值.

参 考 文 献

[1] Drinfeld, V. D., Vladut, S. G., Number of points of an algebraic curves, Func. Anal. 17(1983), 53～54.

[2] Feng, G. L., Rao, T. R. N., Decoding algebraic geometry codes up to the design minimum distance, IEEE Trans. IT－39(1993), 37～45.

[3] Garcia, A., Stichtenoth, H., A tower of Artin-Schreier extensions of function fields attaining the Drinfeld-Vladut bound, Invent. Math. 121(1995), 211～222.

[4] Goppa, V. D., Codes on algebraic curves, Soviet Math. Dokl. 24(1981), 75～91.

[5] ——, Geometry and Codes, Mathematics and its applications, vol. 24, Kluwer Acad. Publ. 1991.

[6] Ihara, Y., Some remarks on the number of rational points of algebraic curves over finite fields, J. Fac. Sc. Univ. Tokyo, 28(1981), 721～724.

[7] Manin, Y. I., What is maximum number of points on a curve over F_2, J. Fac. Sci. Univ. Tokyo, 28(1981), 715～720.

[8] Moreno, C., Algebraic curves over finite fields, Cambridge Univ. Press, 1993.

[9] Serre, J.-P., Sur les nombres des points rationnels d'one courbe algebrique sur un corps fini, Compt. Rendu Azad.

Sci. Paris, 297(1983), 397～401.

[10] Stichtenoch, H., Algebraic Function Fields and Codes, Springer-Verlag, 1993.

[11] Tsfasman, M. A., Vladut, S. G., Algebraic — Geometric Codes, Kluwer, 1991.

[12] Tsfasman, M. A., Vladut, S. F., Zink, T. Modular curves, Shimura curves and Goppa codes better than the Varshamov-Gilbert bound, Math. Nachr. 104(1982), 13～28.

§5.3 大数分解和公开密钥体制

密码和纠错码的功能完全不同，纠错码是希望信息在传输中不产生错误，并且有能力纠正微小的错误，而密码则是人为地把发出的信息加密，使得第三者在收到时不能恢复原始信息，但是真正的收方要能够恢复原文(去密). 在无线通信时代，信息很容易被外人所窃取、篡改和仿造，所以信息加密愈来愈变得重要. 随着通信领域的扩展，信息安全问题已不局限于在国家外交和军事等领域，而广泛深入到人类生活各个领域，特别是经济领域(如电子商务). 传统通信中的许多功能(如签名，身份认证，仲裁，密钥分配和共享等)在电子通信系统中如何实现?如何设计和实现多功能的安全协议? 如何评价一个体系是真正安全的? 这就提出了许多新鲜有趣并且在应用领域非常重要的数学问题. 比如 A 问 B 是否知道某个问题 x 的答案. 在很多情形下 A 的用意并不在于弄清 B 是否知道，而是 A 自己想从 B 那里"套出"问题 x 的答案. 所以 B 应当有一种办法来表明他自己知道问题 x 的答案，但是对答案本身是什

么却不能透露出任何信息，这就是近年来发展的“零知识证明”问题.

目前的通信已不再是两个人之间的联系，而是由众多计算机和通信设备组成的多用户的通信网络. 用户 A 向用户 B 发信息 x，要把它加密为 $E_A(x) = y$，而用户收到 y 之后要能够去密：$D_A(y) = x$. E_A 和 D_A 分别叫作用户 A 的加密和去密的**密钥**，其中 $D_AE_A=I$(恒等运算). 这些密钥只有 A 和 B 之间知道，外人不知道而且破译密钥非常困难. 如果用户有 n 个，则每个用户都要有 $2(n-1)$ 套密钥(他和其余 $n-1$ 个用户相互通信的密钥系统)，整个通信网络共需设计 $2n(n-1)$ 套密钥系统. 当 n 很大时，每个用户需要保存并且不停地更换大量密钥，而密钥的丢失是很严重的问题. 这个问题长期未能得到有效地解决.

1976 年，美国年青的数学家和计算机专家 Diffie 和 Hellman 提出一种新的保密体制，叫作**公开密钥**(public key)**体制**，这种体制的关键是使用所谓“单向”(one way)函数 E 作为加密密钥. 所谓单向函数 E 即指 E 是可逆函数，但是由 E 求出逆函数 $D=E^{-1}$ 是非常困难的. n 个用户 $A_i(1 \leqslant i \leqslant n)$ 当中每个人 A_i 都选取一个单向函数 E_i 作为自己的加密密钥并把它公之于众，但是把 $D_i = E_i^{-1}$ 作为自己的去密密钥保存起来，只有自己知道. n 个用户的加密密钥 $E_i(1 \leqslant i \leqslant n)$ 可以装订成册，就像电话本一样可供任何人查阅. 当 A_i 把信息 x 传给 A_j 时，A_i 在公开的密码本上查到 A_j 的加密密钥 E_j，A_i 把 $y = E_j(x)$ 传给 A_j. A_j 在收到密文 y 之后用只有自己知道的去密密钥 D_j 作用于 y，就得到 A_i 传给他的明文 $x = D_j(y)$. 对于第三者来说，即使他截获到密文 y 并且也知道是 A_i 传给 A_j 的，在查到 E_j 之后很难求出 D_j，所以不能把 y 恢复成明文 x.

对于这种公密体制，每个用户 A_i 只需保存一个密钥 D_i，从而解决了长期困扰人们的密钥保存问题. 此外，这种体制还解决了长期存在的另一个难题，即信息的签名和认证问题. 如果 A_i 发出信

息是向 A_j 要一笔钱，A_j 需要 A_i 在这条信息上签名，以防止别人冒名顶替. 公钥体制可以作这件事：A_i 先用自己的 D_i 作用于 x(签名)，把 $y = D_i(x)$ 传给 A_j. A_j 收到后用查到的 E_i 作用 y，恢复出明文 $x = E_i(y)$，这便认出信息确实来自 A_i，因为只有 A_i 自己知道 D_i. 进而，还可同时进行加密和签名两种功能：A_i 把信息 x 连续作用 D_i(签名)和 E_j(加密)，将 $y = E_jD_i(x)$ 传给 A_j. 而 A_j 将 y 连续作用 D_j(去密)和 E_i(认证)，得到 $E_iD_j(y) = E_iD_jE_jD_i(x) = E_iD_i(x) = x$，从而恢复明文.

公钥体制提出之后，由于解决了密钥保存和数字签名问题，引起通信界和数学界的很大反响. 一时间人们设计出五花八门的公钥具体方案(背包方案，中国剩余定理方案……)，这些方案随后又不断被别人攻击和破掉. 这些方案的核心是设计单向函数 E. 所谓被破掉即指寻找出求 E^{-1}的好算法，以表明采用 E 作为加密密钥是不可靠的. 迄今为止的 20 多年来，只留下两种公钥方案目前被认为是可靠的，这就是大数分解方案和离散对数方案. 它们分别基于大数分解和离散对数问题. 经过数学家的多年研究，目前还没有好的算法.

我们在上节已经说过，目前分解一个 200 位的整数(特别当它是两个差不多大小的不同素数之积时)是非常困难的. 就在公钥体制出现一年之后，1977 年美国麻省理工学院的三位年青学者 Rivest，Shamir 和 Adleman 提出如下的公钥方案，后人称为 RSA 方案. 取定两个大约 100 位左右的不同素数 p 和 q(我们在 § 5.1 说过，找出这样的素数 p 和 q 有多项式算法). 令 $N = pq$. 则它的欧拉函数为 $\varphi(N) = (p-1)(q-1)$. 然后再取两个正整数 e 和 d，使得 $ed \equiv 1 \pmod{\varphi(n)}$ (当 n 很大时，这样的数组 (e,d) 有很多，可供不同人使用). 通信时把明文取为 $0,1,\cdots,N-1$ 当中的整数 x，加密密钥 E 取为 $y \equiv E(x) \equiv x^e \pmod N$，即 y 是 x^e 模 N 的最小非负剩余. 而去密密钥 D 取为 $D(y) \equiv y^d \pmod N$. 由欧拉定理可知 E 和 D 的作用是互逆的，即 $D(y) = x$.

在这种方案中 N 和 e 是公开的. 由 e 求 d 需要满足 $ed \equiv 1 \pmod{\varphi(N)}$，但是这必需知道 $\varphi(N) = (p-1)(q-1)$ 的值. 由于分解 $N = pq$ 非常困难，外人无法知道 $\varphi(N)$，所以无法由 e 求 d.

RSA 方案刺激了人们对大数分解的竞赛热潮. 而人们对这场竞赛却抱着完全不同的心情. 到目前为止，数学家们运用高深的数论知识和思想，创造了许多新的改进算法，尽管有更大的数被分解，但仍是亚指数型算法. 基于此，目前 RSA 方案已经应用到实际领域中. 如果将来发现更好的算法，使大数分解不再困难，这无疑是数学家的胜利，也意味着 RSA 方案的破产.

我们在上节已经介绍了二次筛法和数域筛法的最基本思想(因子基方法)，现在我们介绍椭圆曲线算法进行大数分解的基本思想.

设 $E: y^2 = x^3 + ax + b$ 为椭圆曲线，其中 $a, b \in \mathbf{Z}, 4a^3 + 27b^2 \neq 0$. 则有理点群 $E(\mathbf{Q})$ 上的加法运算定义为 $(x_1, y_1) \oplus (x_2, y_2) = (x_3, y_3)$，其中

$$x_3 = k^2 - x_1 - x_2, y_3 = -y_1 + k(x_1 - x_3).$$

而当 $(x_1, y_1) \neq (x_2, y_2)$ 时，$k = \dfrac{y_2 - y_1}{x_2 - x_1}$. 当 $(x_1, y_1) = (x_2, y_2)$ 时，$k = \dfrac{3x_1^2 + a}{2y_1}$. 如果运算中分数出现 $\dfrac{a}{0}$ $(a \neq 0)$，则得到无穷远点 0 (它是群 $E(\mathbf{Q})$ 的零元素).

现在要把大整数 N 作因子分解. 我们只需求出 N 的一个真因子 d，然后再把 d 和 N/d 分解. 我们随机地取一个椭圆曲线 E (有上述形式的方程)和 E 上一个整数点 P，并且使 $(4a^3 + 27b^2, N) = 1$(如果 $(4a^3 + 27b^2, N) \neq 1$，常常得到 N 的一个真因子 $(4a^3 + 27b^2, N)$. 如果 $(4a^3 + 27b^2, N) = N$，我们换一条椭圆曲线). 我们取一个整数 B(B 的大小以后在分析算法复杂性时再决定)，再取 C 为大于 $N^{\frac{1}{2}} + 1 + 2N^{\frac{1}{4}}$ 的一个整数. 令

$$k = \prod_{p \leqslant B} p^{a_p}, a_p = \left[\frac{\log C}{\log p}\right].$$

例如若 N 是一个 10 位数，$B=20$，则可取 $C=100700$，而 $k=2^{16} \cdot 3^{10} \cdot 5^7 \cdot 7^5 \cdot 11^4 \cdot 13^4 \cdot 17^4 \cdot 19^3$.

如果 N 是合成数，则它必有素因子 $p \leqslant N^{\frac{1}{2}}$. 由于 $(4a^3 + 27b^2, N) = 1$，可知 E 模 p 是有限域 $\boldsymbol{F}_p$ 上的一条椭圆曲线，记为 E_p. 由 Hasse-Weil 定理，E_p 在 $\boldsymbol{F}_p$ 上的点群 $E_p(\boldsymbol{F}_p)$ 的阶 N_p 满足 $N_p \leqslant p+1+2\sqrt{p} \leqslant N^{\frac{1}{2}}+1+2N^{\frac{1}{4}} \leqslant C$. 取 B 足够大，则有很大的概率使得 N_p 的素因子均不超过 B. 这时由 k 的定义即知 $N_p | k$. 即对 E_p 上的整数点 P(坐标 mod p)，P 的 k 倍一定是 O(mod p).

于是我们依次计算下列点的 mod N 坐标：

P, $[2]P$, $[2^2]P$, …, $[2^{a_2}P]$, $[2^{a_2} \cdot 3]P$, $[2^{a_2} \cdot 3^2]P$, …, $[2^{a_2} \cdot 3^{a_3}]P$, $[2^{a_2} \cdot 3^{a_3} \cdot 5]P$, …, $[2^{a_2} \cdot 3^{a_3} \cdot 5^{a_5}]P$, …, $[k]P$.

由于加法运算表达式中出现分数，所以模 N 运算时分母与 N 互素才可进行下去. 但是这些点当中有很大的概率使得 mod p 之后是无穷远点(其中 p 是 N 的一个素因子). 所以在计算上述点列的坐标时，每次模 N 运算对出现的分母，需要求分母与 N 的最大公因子，有很大概率使得这个最大公因子 d 大于 1. 如果 $d<N$，就得到 N 的一个真因子. 如果 $d=N$，则换曲线 E 上另一个整点 P'，或者换一条椭圆曲线和其上一个整点继续作下去，一直到找出 N 的一个真因子为止. 综上所述，这是一个概率算法. 通过对算法复杂性的分析，得出取 $B \approx \exp(\sqrt{\log N \log\log N})$ 时算法最佳，并且该算法的计算量为 $\exp\sqrt{(1+\varepsilon)\log N \cdot \log\log N}$，即为亚指数型算法.

现在谈离散对数问题在信息安全中的应用. 我们在 §5.1 中讲过，所谓离散对数问题就是对一个有限阿贝尔群 G 和 G 中两个元素 g 和 $h=g^x$，由 g 和 h 求整数 $x(1 \leqslant x \leqslant |G|)$. 目前对许多具

体的群 G,这个问题只有亚指数型算法,所以人们用它来实现信息安全的许多功能.例如

(1) 密钥交换.用户 A 和 B 之间的密钥常常需要更换,所以过一段时间要商定一个新的密钥,而这个密钥不能被外人知道.1977 年 Diffie 和 Hellman 给出如下的方案:选取一个有限阿贝尔群 G 和 G 中一个阶数很大的元素 g.用户 A 随机取一个整数 x_A,$1 \leqslant x_A \leqslant |G|-1$,计算 g^{x_A},把它传给 B.而 B 也随机地取一个整数 x_B,$1 \leqslant x_B \leqslant |G|-1$,计算 g^{x_B}并传给 A.A 收到 g^{x_B}之后算出 $(g^{x_B})^{x_A}$,B 收到 g^{x_A}之后算出 $(g^{x_A})^{x_B}$.于是 A 和 B 用这个公共值 $g^{x_A x_B}$作为密钥.外人只知道 G, g, g^{x_A}和 g^{x_B}.由于离散对数问题是困难的,由这些资料无法算出 x_A 或 x_B,所以无法得到密钥 $g^{x_A x_B}$.

(2) 1985 年,EIGamal 用离散对数给出如下的公钥方案.将 G, g 和 $h=g^x$ 公开,但是 x 保密.设 A 要把信息 $m(\in G)$ 传给 B,A 随机地选取 k,$1 \leqslant k \leqslant |G|-1$,计算 $a=g^k$ 和 $b=h^k m$,把数据组 (a,b) 传给 B.B 计算 $ba^{-x}=h^k m g^{-kx}=m$,恢复信息 m.密钥 x 只有 B 知道.第三者由 g,h,a 和 b 计算 m 很困难.

(3) EIGamal 于 1985 年用离散对数还构作了数字签名方案.假设 B 要把信息 $m \in \mathbf{Z}/|G|\mathbf{Z}$ 传给 A,二人事先约定一个一一对应映射 $f: G \to \mathbf{Z}/|G|\mathbf{Z}$.公开 G, g 和 $h=g^x$,而 x 为 B 自己的密钥.方法是:B 随机地选取 k,$1 \leqslant k \leqslant |G|-1$,$(k,|G|)=1$.计算 $a=g^k$,并且求出同余方程 $m \equiv xf(a)+bk(\bmod |G|)$ 的一个解 $b \in \mathbf{Z}/|G|\mathbf{Z}$.然后把 (a,b)(B 的签名)和信息 m 一起传给 A.A 计算 $h^{f(a)}a^b$,验证它等于 g^m,这就确信信息来自 B,因为外人不知 x,从而无法算出 b.

在目前的算法中,G 取成有限域 $\boldsymbol{F}_q$ 的乘法循环群 $\boldsymbol{F}_q^{\times}$,而 g 取成它的一个生成元.1987 年,Koblitz 建议用 $\boldsymbol{F}_q$ 上椭圆曲线 E 的点群 $E(\boldsymbol{F}q)$,而 g 取成其中的大阶元素.1988 年 Williams 等人建议用虚二次域的理想类群.理想类和二元二次型有一一对应关系,而理想类的乘法对应于高斯对于二元二次型的合成运算,狄德金

把二元二次型理论改造成理想论，具有重要的理论意义. 现在从实际需要又把理想乘法具体用高斯的二元二次型合成运算来实现，充分体现出理论研究和应用领域不同的美学标准. 1992 年～1996 年，Adleman 和黄铭德把 G 取成有限域上超椭圆曲线 $y^2 = f(x)$ 的除子类群和高维阿贝尔簇的群结构.

参 考 文 献

[1] L. Adleman and M. Huang, Primality testing and Abelian varieties over finite fields, LN in Math. 1512, Springer-Verlag, 1992.

[2] N. Koblitz, A Course in Number Theory and Cryptography, Springer-Verlag, 1987.

[3] L. Adleman, R. Rivest and A. Shamir, A method for obtaining digital signatures and public-key cryptosystems. Com. ACM, 21(1978), 120～126.

[4] W. Diffie, and M. Hellman, New directions in cryptography, IEEE Trans. IT-22(1976), 644～654.

[5] D. Kahn, The Codebreakers, the story of secret writing. Macmillam, 1971.

[6] H. W. Lenstra, Factoring integers with elliptic curves, Ann. of Math. 126(1987), 649～673.

[7] A. K. Lenstra and H. W. Lenstra (eds.), The development of the number field sieve, LN in Math. 1554, Springer-Verlag, 1994.

[8] N. Koblitz, Elliptic curve cryptosystems, Math. Comp. 48

(1987),203～209.
[9] A. Odlyzko, Discrete logarithms in finite fields and their cryptographic significance, Advances in Cryptology, Proceedings of Eurocrypt 84, Springer-Verlag 1985, 224～314.
[10] T. ElGamal, A public key cryptosystem and a signature scheme based on discrete logarithms. IEEE Trans. IT-31 (1985)469～472.
[11] A. J. Menezes, Elliptic Curve Public Key Cryptosystems. Kluwer Academic Pub. 1993.
[12] A. J. Menezes, P. C. van Oorschot and S. A. Vanstone, Handbook of Applied Cryptography, CRC Press, 1996.
[13] I. Blake, G. Seroussi and N. Smart, Elliptic Curves in Cryptography, Cambridge Univ. Press, 1998.
[14] K. Nyberg and R. A. Rueppel, Message recovery for signature schemes based on the discrete logarithm problem, Designs, Codes and Cryptography, 7(1996), 61～81.
[15] H. -G. Rück, On the discrete logarithm problem in the divisor class group of curves. Math. Comp. 68(1999), 805～806.
[16] B. Schneier, Applied Cryptography, John Wiley and Sons, 1996.
[17] S. A. Vanstone and R. J. Zuccherato, Elliptic curve cryptosystems using curves of smooth order over the ring Zn, IEEE Trans. IT-43(1997), 1231～1237.

§ 5.4 自守表示和通信网络
(Ramanujan 图)

本节谈如何设计一个好的通信网络. 一个通信网络就是一个图. 以有线电话系统为例,假设有 n 个电话用户,表示成 n 个顶点. 如果用户 A 和 B 可以直接通话(有电话线相连),便在顶点 A 和 B 之间引一条线,叫作以 A 和 B 为两端的一条边,表示成 $\overline{AB}$. 这样由 n 个顶点组成的顶点集合 V 与一些边组成的集合 E,就构成一个(无向)图 $G=(V,E)$,并且以后假设图是简单的,即边的两端是不同顶点,并且任两个不同顶点之间至多有一条边.

首尾相接的一些边 $\overline{A_1A_2},\overline{A_2A_3},\cdots,\overline{A_nA_{n+1}}$ 组成图 G 的一条路,边数 n 叫作路长,这时,A_1 的信息通过这条路经过 n 步传递可达到 A_{n+1}. 如果图中任意两个不同顶点均有路相连,称 G 为连通图. 如果顶点 A 向外一共引出 k 条边,k 叫顶点 A 的次数,表示成 $\deg A$,而这 k 条边的另一端叫作 A 的 k 个邻居. 如果每个顶点的次数均为 k,称 G 为 k 次正则图. 在连通图 G 中,两点之间的最短路长叫作这两点的距离,所有点之间距离的最大值叫作图 G 的直径,表示成 $d=d(G)$.

设 G 是 n 个用户的通信网络(n 个顶点的图). 一个好的通信网络希望信息传输很快. 任何用户的信息经过 $d=d(G)$ 步都能传给网络中所有用户,并且 d 是满足此条件的最小值. 所以我们希望直径 d 愈小愈好. 显然 d 达到最小值 1 当且仅当任意两个不同顶点都有边相连,这样的 n 点图叫作完全图,表示成 K_n. K_n 共有 $n(n-1)/2$ 条边. 当 n 很大时,这种网络很不经济. 所以我们设每

个点的次数均不超过 k. 或者考虑比较简单的情形，即每个点的次数均为 $k(k < n)$，即 G 为 k 次正则图的情形. 我们要问：n 点 k 次正则图的最小直径是多少？

通信中衡量网络效率的另一个参量叫作放大系数. 设 X 是图 G 顶点集合的子集合. 如果某个时刻 X 中所有用户均知道某个信息，然后 X 中每个用户把信息传给他的所有邻居，则下一时刻知道此信息的新用户集合为

$$\partial X = \{v \in V \mid v \notin X, \text{存在 } x \in X \text{ 使 } v \text{ 是 } x \text{ 的邻居}\}.$$

网络 G 的**放大系数**定义为

$$C = C(G) = \min\left\{\frac{|\partial X|}{|X|} \,\Big|\, X \subset V, |X| \leqslant \frac{1}{2}|V|\right\}.$$

例如当 n 为偶数时 $c(K_n) = 1$，而 n 为奇数时，$c(K_n) = \frac{n+1}{n-1}$. 如果 G 不是完全图，则 $0 \leqslant c \leqslant 1$. 当网络很大时决定放大系数是相当困难的. 那么，n 个顶点的 k 次正则图，其放大系数最大值是多少？如何具体构作出 $C(G)$ 很大的 n 顶点 k 次正则图？更一般地，如何构作出一个 k 次正则图序列，使得顶点个数趋于无穷，并且这些图的放大系数都很大？

构作这种理想的通信网络族的第一个系统方法是由菲尔兹奖获得者 Margulis 于 1975 年给出的，他利用了无限群表示论中的 Kazhdan 性质 T. 局部紧群 Γ 叫作具有性质 T，是指 Γ 的平凡表示孤立于其它酉表示，而衡量孤立程度是用所谓 Kazhdan 常数 k. 注意表示集合上的拓朴类似于通常的"紧开"拓朴，但是紧集在这里改用群 Γ 的某个生成元集合. 所以平凡表示与其它酉表示的"距离"K 依赖于生成元集合的选取. Margulis 使用群 $\Gamma = SL_2(\mathbf{Z}) \times \mathbf{Z}^2$（半直积）和 Γ 的四个元素组成的标准生成元，得到无穷多个 5 次正则双分图，其放大系数均为 $K^2/2$. 这里，一个图叫做是**双分**的，是指顶点集合 V 可以分拆成两个部分 V_1 和 V_2，使得图中所有边均是一端在 V_1，另一端在 V_2. 1981 年 Gabber 和 Galil 改用 Γ 的另外四个元素作为生成元，得到另一个 5 次正则图族，其常数 K

可以明确计算出来. 此后,又陆续构作出放大系数不断改进的通信网络.

差不多与此同时,人们发现一个好的通信网络可以用一个代数不变量来控制. 这个不变量是图的**次根**. 对于一个 n 顶点图 G,设顶点集合为 $V=\{v_1,v_2,\cdots,v_n\}$,如下构作 n 阶方阵 $A=A(G)=(a_{ij})$,其中

$$a_{ij}=\begin{cases}1, & \text{若图 } G \text{ 中有边 } \overline{v_iv_j},\\ 0, & \text{否则}.\end{cases}$$

这是一个 n 阶实对称方阵,所以 A 的 n 个特征根都是实数. 将这些特征根从大到小依次列出:$\lambda_0\geqslant\lambda_1\geqslant\cdots\geqslant\lambda_{n-1}$. 若 G 是 k 次正则连通图,则当 G 是双分图时 $\lambda_0=k,\lambda_{n-1}=-k$,其它特征根的绝对值均小于 k. 若 G 不是双分图,则 $\lambda_0=k$,其它特征根的绝对值均小于 k. 特征根绝对值 $|\lambda_i|$ 当中小于 k 的那些的最大值叫作图 G 的**次根**,表示成 $\lambda=\lambda(G)$. 用方阵 $A=A(G)$(叫作图 G 的**关联方阵**)的各种代数特性来刻划和研究图 G 的各种图论特性,是图论的一个专门方向,叫作**代数图论**. 例如,对每个正整数 l,令 $A^l=(a_{ij}^{(l)})$,则 $a_{ij}^{(l)}$ 就是图 G 中以顶点 v_i 和 v_j 为两端长为 l 的路的个数. 以 D 表示对角方阵,其主对角线元素为 $(\deg v_1,\cdots,\deg v_n)$,则方阵 $L=L(G)=D-A(G)$ 是正定对称阵,特征根为 $0\leqslant l_0\leqslant l_1\leqslant\cdots\leqslant l_{n-1}$. 如果 G 是 k 次正则连通图,则 $D=kI_n$ 而 $l_i=k-\lambda_i$. 于是 $l_0=0$ 而其余 l_i 均为正实数.

人们陆续发现通信网络的特性和图的特征根之间有如下的关系:

(1) 设 G 是连通图并且每个顶点次数均不超过 k,c 表示 G 的放大系数. 则

$$c\geqslant\frac{2l_1}{k+2l_1}\text{(Tanner,1984)},$$

$$l_1\geqslant\frac{c^2}{4+2c^2}\text{(Alon-Milman,1985)}.$$

(2)（1989,金芳蓉） 若 G 为 n 个顶点的 k 次正则连通图，则

$$d(G) \leqslant \frac{\log n - 1}{\log k/\lambda(G)}.$$

对于 k 次正则连通图，$\lambda(G) \geqslant k - k_1$. 由上述结果可知对 $\lambda(G)$ 小的连通正则图，有大的放大系数 C 和小的直径，这正是好的通信网络. 例如对 n 个顶点的完全图 K_n，关联方阵 $A(K_n)$ 是 n 阶方阵，主对角阵上元素为 0，其余元素均为 1. 它的特征根为 $n-1$ 和 -1($n-1$ 重根)，于是 $\lambda(G) = 1$. 这虽然是好的通信网络，但很不经济（边数太多）. 于是人们问：对于固定的正整数 k，一个 k 次正则连通图 G 的次根 $\lambda(G)$ 能小到何种程度？我们有如下的渐近性下界

定理 设 $\{G_n\}$ 是任意一族 k 次正则连通图，G_n 的顶点数趋于无穷，则 $\lim\lambda(G_n) \geqslant 2\sqrt{k-1}$.

这个结果最早是 Mckay 于 1981 年证明的，后来人们又给出许多种证明方法和不同的理解方式. 一种是随机的观点：在所有 k 次正则连通图当中，绝大多数图的次根都在 $2\sqrt{k-1}$ 附近. 另一种是覆盖的观点：所有 k 次正则连通有限图 G 都存在一个通用的覆盖，即是 k 次正则无限树 T_k. T_k 有无限多个顶点，每个顶点有 k 个邻居并且没有回路. 对每个 k 次正则连通有限图 G 均存在 T_k 的自同构群的一个子群 N，使得 T_k 对于 N 的"商图"就是 G. 考虑定义在 T_k 的无限个顶点上的实函数全体（记为 S），这是实线性空间. 可以定义其上的一个线性算子 $\mathscr{A}$：对每个 $f \in S$，$\mathscr{A}(f)$ 在顶点 v 处的取值为 f 在 v 的 k 个邻居上取值之和，这正是有限图关联方阵的一种（在无限图上的）模拟. 对无限树 T_k 赋以一定拓朴结构，由此使 S 为希尔伯特空间，而 $\mathscr{A}$ 是 S 的有界线性自伴算子. 通过发展 T_k 的自同构群在空间 S 上的表示理论，得到算子 $\mathscr{A}$ 的

次根为 $2\sqrt{k-1}$. 然后经过"商"运算可知 k 次正则连通有限图族的次根极限值$\geqslant 2\sqrt{k-1}$. 图覆盖方法和上述表示论工具在历史上曾经用来研究无限群的许多深刻性质. 例如证明 Schreier 定理(自由群的子群是自由群)和 Ihara 定理($SL_2(Q_P)$的无扭子群为自由群)等,详见 Serre《树》一书. 第三种观点是把图看成是紧黎曼流形的离散模拟,紧黎曼流形上的拉普拉斯算子对应于图的顶点集合上的函数,它的矩阵表示就是 $L=D-A$(从而 L 也叫作图的拉普拉斯算子). 将微分几何与调和分析的各种结果离散化,可以得到用图 G 的谱 $\lambda_0\leqslant\lambda_1\leqslant\cdots\leqslant\lambda_{n-1}$来估计直径 $d(G)$ 和其它图参数的各种结果.

基于上述定理,Lubotzky,Phillips 和 Sarnak 三人于 1988 年引入下面的概念.

定义 k 次正则连通图 G 叫作 Ramanujan 图,是指 $\lambda(G)\leqslant 2\sqrt{k-1}$.

例如对 $k=2$,2 次正则连通图即是 n 个点连成的一个回路 C_n. C_n 的大根为 2,所以次根 $\lambda(C_n)<2=2\sqrt{k-1}$. 这表明每个 2 次正则连通图均是 Ramanujan 图. 当 $k\geqslant 3$ 时,Ramanujan 图的构作方法是用群来构作 Cayley 图.

设 M 是有限群,S 为 M 的一个子集合,$1\notin S$ 并且 $S=S^{-1}$. 如下构作图 $G=G(M,S)$:G 的顶点集合为 M,而对每个顶点 $a\in M$,它的邻居为 $ag(g\in S)$. 这个图叫作群 M 对集合 S 的 Cayley 图. 不难看出这是 k 次正则图,其中 $k=|S|$. 并且 G 是连通的当且仅当集合 S 生成群 G. 设 $M=\{v_1,\cdots,v_n\}$,则图 G 的关联方阵为 $A=A(G)=(a_{ij})$, 其中

$$a_{ij}=\begin{cases}1, & 若\ v_i^{-1}v_j\in S,\\ 0, & 否则.\end{cases}$$

这正是 Frobenius 建立群表示论的初期所研究的群矩阵，可以用有限群 M 的复表示理论来计算方阵 $A(G)$ 的特征根. 特别当 M 是 n 阶有限阿贝尔群时，$A(G)$ 的 n 个特征根为

$$\lambda(\chi) = \sum_{g \in S} \chi(g),$$

其中 χ 过群 M 的 n 个特征. 当 χ 为平凡特征(取值恒为 1)时，$\lambda(\chi) = |S|$ 即是图 G 的大根. 而次根的估计就归结于上述特征和的估计. 有限域上代数曲线的韦依定理，有限域上高维代数簇的 Deligne 定理以及自守表示 L-函数的零点结果都可以给出上述特征和绝对值上界估计的许多深刻结果. 利用这些结果，李文卿(1992 年)采用各种有限阿贝尔群 M 和适当的子集 S，构作了一系列 Ramanujan 图. Terras 等人用有限非阿贝尔群 $GL_2(\boldsymbol{F}_q)$ 构作出 $q+1$ 次 Ramanujan 图.

基于通信网络背景，人们提出如下的问题：

对于每个正整数 $k \geqslant 2$，是否存在无限个 k 次 Ramanujan 图族 $\{G_m\}$，使得当 $m \to +\infty$ 时，G_m 的顶点数趋于 $+\infty$？

当 $k=2$ 时，$\{C_n | n = 3,4,5,\cdots\}$ 就是这样的图族. 第一批非平凡的例子是由 Lubotzky, Phillips 和 Sarnak 用无限群的自守表示理论给出的.

考虑局部域 $\boldsymbol{Q}_p$ (p 为素数). 群 $GL_2(\boldsymbol{Q}_p)$ 的每个元素 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 对应于 $\boldsymbol{Q}_p^2$ 中一个格 $L = \boldsymbol{Z}_p e_1 \oplus \boldsymbol{Z}_p e_2$，其中 $e_1 = \begin{pmatrix} a \\ c \end{pmatrix}$，$e_2 = \begin{pmatrix} b \\ d \end{pmatrix}$. $GL_2(\boldsymbol{Q}_P)$ 中元素 g_1 和 g_2 对应同一个格当且仅当 $g_1^{-1} g_2 \in GL_2(\boldsymbol{Z}_p)$. 所以可把商群 $GL_2(\boldsymbol{Q}_p)/GL_2(\boldsymbol{Z}_p)$ 等同于 $\boldsymbol{Q}_p^2$ 中的格集合. 进而，格 L_1 和 L_2 相似是指存在 $x \in \boldsymbol{Q}_p^x$ 使得 $L_1 = xL_2$. 于是群

$$M = GL_2(\boldsymbol{Q}_p)/GL_2(\boldsymbol{Z}_p)\boldsymbol{Q}_p^x$$

等同于 $\boldsymbol{Q}_p^2$ 上格的相似等价类集合. 如下构作一个图 T：T 的顶点集合为 M，而顶点 v_1 和 v_2 有边相连当且仅当相似等价类 v_1 和 v_2

中分别有格 L_1 和 L_2，使得 L_2 为 L_1 的子格，并且 $[L_1 : L_2] = p$（注意若 $[L_1 : L_2] = p$，则 $[L_2 : pL_1] = p$，从而这定义出边 $\overline{v_1v_2}$）．不难证明格 $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 的相似类共有 $p+1$ 个邻居 $g \cdot h$，其中 $h = \begin{pmatrix} pu \\ 01 \end{pmatrix}$ $(0 \leqslant u \leqslant p-1)$ 和 $\begin{pmatrix} 1 & o \\ o & p \end{pmatrix}$．由此可知 T 是 $p+1$ 次无限正则树 T_{p+1}．$GL_2(\boldsymbol{Q}_p)$ 的适当子群 Γ 作用在 T 上，其商图 $X = \Gamma \backslash T = \Gamma \backslash GL_2(\boldsymbol{Q}_p) / GL_2(\boldsymbol{Z}_p)\boldsymbol{Q}_p^x$ 就是 $p+1$ 次正则有限图．以 S 表示图 X 顶点集合上的函数全体，每个函数 $f \in S$ 可以看成是 $GL_2(\boldsymbol{Z}_p)\boldsymbol{Q}^x$ 不变的 $GL_2(\boldsymbol{Q}_p)$ 上的自守函数．用 Adèle 语言将上述群作一些变换，则 f 成为 $GL_2(\boldsymbol{A}_Q)$ 上的权 2 尖点自守形式．图 X 的关联矩阵 $A(X)$ 看成空间 S 上的线性算子，在上面的变换下它恰好是权 2 自守形式空间上的 Hecke 算子 T_p．于是利用 $GL_2(\boldsymbol{A}_Q)$ 的表示理论：对于平凡表示，Hecke 算子 T_p 的特征函数即为常值函数，而特征根 $p+1$ 即是图 X 的大根．对于非平凡表示，特征函数为权 2 尖点自守形式，根据 Deligne 证明的 Ramanujan-Petersson 猜想，这些自守形式作为 T_p 的特征函数，其特征根 λ 满足 $|\lambda| = 2\sqrt{p} \leqslant 2\sqrt{k-1}$．即上面作出的 $p+1$ 次正则图 X 的次根 $\lambda(X) \leqslant 2\sqrt{p}$，从而为 Ramanujan 图．取不同的群 Γ 可给出 $p+1$ 次正则 Ramanujan 图族．换句话说，对每个素数 p，存在着 $p+1$ 次正则 Ramanujan 图族．1988 年，Drinfeld 对于函数域情形证明了 GL_2 上关于 Ramanujan 猜想的模拟．利用这个结果，Morgenstern 于 1994 年对每个素数幂 q，证明存在 $q+1$ 次正则 Ramanujan 图族．对于其它 k 值，k 次正则 Ramanujan 图族的存在性问题至今未能解决．

与 Ramanujan 图有关的是近年来人们关注的另一个有趣的图论问题：图的 zeta 函数．它最早由 Ihara 于 1966 年给出定义，用

来研究 p-adic 域上 2 阶射影群的离散子群. 1990 年前后 Hashimoto 把它推广成 Artin L-函数在图上的一种类比，用来研究 p-adic 群的表示. 与此同时 Sunada 研究图上的 zeta 函数，看作是紧黎曼流形拉普拉斯算子的谱构作的 Selberg zeta 函数的离散化形式. 在后来对图 zeta 函数的研究中，H. Bass(1992)使用了更多的代数工具，Venkov 和 Nikitin 则采用 Selberg 迹公式，Stark 和 Terras(1996)使用较初等的组合方法，并且给出一些有趣的例子以显示与数域 zeta 函数的相似. Foata 和 Zeilberger(1999)研究了图上 zeta 函数与符号语言学之间的联系.

现在介绍图的 zeta 函数定义和基本性质. 设 X 为连通(简单)图. 若$\overline{v_1v_2},\overline{v_2v_3},\cdots,\overline{v_{m-1}v_m},\overline{v_mv_1}$均是图中的边，称它们是一条长为 m 的**回路**，并且表示成 $C=(v_1,v_2,\cdots,v_m)$. 以 $l(c)$ 表示它的长度 m. 回路 c 叫作**无返回**的，是指对每个 $i(2\leqslant i\leqslant m)$，$v_{i-1}\neq v_{i+1}$(令 $v_{m+1}=v_1$). 例如下图是长为 9 的无返回的回路. 去掉"尾巴"成为**无尾**的回路 (v_3,v_4,v_5,v_6,v_7). 无返回和无尾的回路叫作**约化回路**. 两个长为 l 的约化回路 $(v_1,v_2,\cdots,v_l)$ 和 $(u_1,u_2,\cdots,u_l)$ 叫作等价，是指存在整数 d，使得 $v_i=u_{i+d}(1\leqslant i\leqslant l$，顶点的下标 mod l). 以$[c]$表示约化回路 C 的等价类. 对两个回路 $C=(v_1,\cdots,v_m)$ 和 $C'=(u_1,\cdots,u_n)$. 如果 $v_1=u_1$，则回路 $(v_1,\cdots,v_m,v_1=u_1,u_2,\cdots,u_n)$ 叫作 C 和 C' 的乘积，表示成 $C\cdot C'$. 约化回路 C 叫作**本原**的，则指不存在回路 C' 和整数 $t\geqslant 2$，使得 $C=(C')^t$.

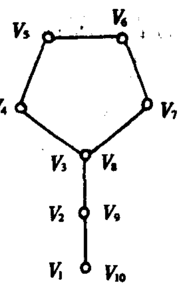


图 16

定义(Ihara)　对每个连通图 X，它的 zeta 函数定义为

$$Z_X(u)=\prod_{[C]}(1-u^{l(C)})^{-1},$$

其中$[C]$过 X 的所有本原回路等价类.

在代数拓朴中，连通图 X 看作为一维复形，对任意顶点 v，以 v 为起点和终点的所有回路（即闭链）的同伦类对于上面定义的运算形成群 $\Gamma=\pi_1(X,v)$，叫 X 的基本群. 这是秩为 r 的自由群，其中 $r=$图 X 的边数$-$图 X 的点数$+1$. 不难看出，约化回路等价类 $[C]$ 和群 Γ 中元素 C 的共轭类$\{C\}$是一一对应的. 将回路 C 看成是群 $\pi_1(X,v)$ 中元素时，若 $C\neq 1$，则 C 在此群中的中心化子为循环群. 而 C 是本原回路当且仅当 C 的中心化子是由 C 生成的循环群. 由代数拓朴知道，一维复形 X 有通用覆盖树 $\widetilde{X}$（若 X 为 $(q+1)$- 正则图，则 $\widetilde{X}$ 为$(q+1)$- 正则树）. 基本群 $\Gamma=\pi_1(X,v)$ 自然作用于 $\widetilde{X}$ 上，并且 X 同构于商图 $\widetilde{X}/\Gamma$. 于是图 X 的 zeta 函数也可表示为

$$Z_x(u)=\prod_{\{C\}}(1-u^{l(C)})^{-1},$$

其中 $\{C\}$ 过基本群 Γ 的所有本原共轭类$(c\neq 1)$，但是 $\{c\}$ 和$\{c^{-1}\}$为不同共轭类.

记 $V=\{v_1,\cdots,v_n\}$ 表示图 X 的顶点集合，$\deg(v_j)=q_j+1$ 是顶点 v_j 的次数. $A=A(X)$ 为前面定义的图 X 关联方阵，$Q=Q(X)$ 为 n 阶对角阵，对角元素为 $\{q_1,q_2,\cdots,q_n\}$. 则图 X 的边数为 $m=\frac{1}{2}\sum_{j=1}^{n}(q_j-1)$，于是（$T_rM$ 表示方阵 M 的迹）

$$r-1=m-n=\frac{1}{2}\sum_{j=1}^{n}(q_j-1)=\frac{1}{2}T_r(Q-I_n).$$

图 X 的 zeta 函数的第一个基本结果为

（1）（Ihara）$Z_X(u)^{-1}=(1-u^2)^{r-1}\det(I-Au+Qu^2)$，并且它是常系数为 1 的整系数多项式；

（2）以 N_m 表示 X 中长为 m 的约化回路的个数，则 $u\frac{\mathrm{d}}{\mathrm{du}}(\log Z_X(u))=\sum_{m\geq 1}N_m u^m$，即　$Z_X(u)=\exp\left(\sum_{m\geq 1}\frac{N_m}{m}u^m\right)$.（注意此公式与有限域上曲线 zeta 函数定义的相似之处.）

例 设 $X=C_n$ 是 n 个顶点组成的一个回路 $(n\geqslant 3)$，则本原回路（等价类）只有两个（方向不同），长度均为 n，因此 $Z_X(u)=(1-u^n)^{-2}$. 或者用结果(2)，长为 m 的约化回路个数

$$N_m=\begin{cases}2n, & \text{若 } m\mid n,\\ 0, & \text{否则}.\end{cases}$$

于是

$$Z_X(u)=\exp\left(2\sum_{i\geqslant 1}\frac{n}{ni}u^{ni}\right)=\exp(-2\cdot\log(1-u^n))$$
$$=(1-u^n)^{-2}.$$

也可用结果(1)，$Z_X(u)^{-1}=\det((1+u^2)I_n-Au)=(1-u^n)^2$.

由结果(1)知：对于连通 $(q+1)$ 次正则图 X，$Z_X(u)^{-1}=(1-u^2)^{r-1}\det((1+qu^2)I_n-Au)$，其中 $r-1=\dfrac{n(q-1)}{2}$.

关于 zeta 函数和 Ramanujan 图的联系为：

(3) 设 X 是 $(q+1)$ 次连通正则图. 则 X 是 Ramanujan 图当且仅当 $Z_X(u)$ 满足如下"黎曼条件"：

如果 $Z_X(q^{-s})=0$ 并且 $0<\mathrm{Re}(s)<1$，则 $\mathrm{Re}(s)=\dfrac{1}{2}$.

进而，图的 zeta 函数满足如下一些函数方程.

(4) 设 X 为 $(q+1)$ 次连通正则图. 令

$$\Lambda_X(u)=(1-u^2)^{\frac{n}{2}+r-1}(1-q^2u^2)^{n/2}Z_X(u)\quad (n \text{ 为 } X \text{ 的顶点个数}),$$

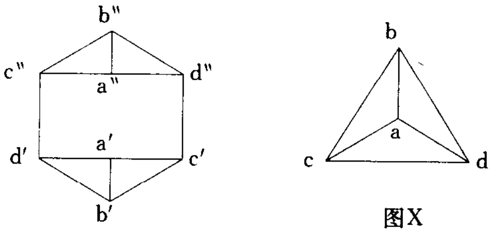
$$\xi_x^{(u)}=(1+u)^{r-1}(1-u)^{r-1+n}(1-qu)^nZ_X(u),$$
$$F_X(u)=(1-u^2)^{r-1}(1+qu^2)^nZ_X(u).$$

则 $\Lambda_X(u)=(-1)^n\Lambda_X\left(\dfrac{1}{qu}\right)$，$\xi_X(u)=\xi_X\left(\dfrac{1}{qu}\right)$，$F_X(u)=F_X\left(\dfrac{1}{qu}\right)$.

接下来讨论图的覆盖与 zeta 函数之间的联系. 设 X 和 Y 是连通（有限）图. 顶点集合上的映射 $\pi:Y\to X$ 叫作图的**同态**，是指 Y 中相邻顶点映成 X 中相邻顶点. 图的满同志 π 叫作 Y 对 X 的一个**覆盖**. 这个覆盖 π 叫作是**不分歧**的，是指若 $\pi(y)=x$，则顶点 y

的邻居和顶点 x 的邻居在 π 之下一一对应. 如果 X 是 $(q+1)$ 次正则图, Y 为 X 的不分歧覆盖, 则 Y 也是 $(q+1)$ 次正则图, 并且对 X 的每个顶点 x, 它的原像 $\pi^{-1}(x)$ 均为 Y 中同样多个(d 个)顶点. 这时也称 Y 是 X 的 d 叶覆盖. 一个 d 叶覆盖 $\pi: Y \to X$ 叫作伽罗华覆盖, 是指 Y 有 d 个自同构 σ, 使得对每个顶点 $y \in Y$, $\pi(\sigma(y)) = \pi(y)$. 这 d 个自同构形成群, 表示成 $\mathrm{Gal}(Y/X)$, 叫 π 的覆盖群, 这些概念来源于黎曼面的覆盖理论, 而图的伽罗华覆盖类比于数域(或函数域)的伽罗华扩张.

(5)设 X 和 Y 均为 $(q+1)$ 次连通正则图, 若 Y 是 X 的覆盖, 则多项式 $Z_X(u)^{-1}$ 为多项式 $Z_Y(u)^{-1}$ 的因子. (即 $Z_Y(u)^{-1}/Z_X(u)^{-1} \in \mathbf{Z}[u]$, 这类比于至今未完全证明的 Artin 猜想: 若 L/K 为数域的扩张, 则 $S_L(s)/S_K(s)$ 是全纯函数.)



图Y　　图X

图 17

例(Stark—Terras)　图 Y 是图 X 的伽罗华覆盖. 覆盖 π 把 Y 中两个顶点 V' 和 V'' 映成 X 中顶点 $v(v = a, b, c, d)$. 这是 2 叶覆盖, 而 $\mathrm{Gal}(Y/X) = \{I, \sigma\}$, 其中 I 为 Y 的恒等自同构, 而 σ 把 Y 的顶点 v' 和 v'' 互换 $(v = a, b, c, d)$. 用结果(1)可算出

$$Z_X(u)^{-1} = (1-u^2)^2 p(u),$$

$$Z_Y(u)^{-1} = (1-u^2)^4(1-u+2u^2)(1+u+2u^2)(1-u^2+4u^4)p(u),$$

其中 $p(u) = (1-u)(1-2u)(1+u+2u^2)^3$. 事实上, 可以构作出图 Y 的 L-函数 $L_Y(u, \chi)$, 其中 χ 是群 $\mathrm{Gal}(Y/X) = \{I, \sigma\}$ 的非

平凡特征,使得 $Z_Y(u)^{-1}=Z_X(u)^{-1}\cdot L_Y(u,\chi)$,就像对数域二次扩张 L/K,$\zeta_L(s)=\zeta_K(s)\cdot\zeta_L(s,\chi)$,其中 $\zeta_L(s,\chi)$ 是 Artin L-函数.

事实上,更一般地,若 $\pi:Y\to X$ 为连通图的不分歧覆盖(Y 和 X 不必正则),则多项式 $Z_X(u)^{-1}$ 是 $Z_Y(u)^{-1}$ 的因子. Stark-Terras 给出一些有趣的例子,以显示与数论 zeta 函数许多结果的相似之处. 如上所述,图 zeta 函数的研究来源于数学许多领域. 能否用它得到图论本身的新结果,特别地,能否用来构作新的 Ramanujan 图族,是一个有意义的课题.

参 考 文 献

关于 Ramanujan 图.

[1] Angel, J., Celniker, N., Poulos, S., Terras, A. 等, Special function on finite upper half planes, Contemp. Math., 138 (1992), 1~26.

[2] Bien, F., Constructions of telephone networks by group representations, Notices of Amer. Math. Soc., 36(1989), 5~22.

[3] Chiu, P., Cubic Ramanujan graphs, Combinatorica, 12 (1992), 275~285.

[4] Chung, F. R. K., Diameters and eigenvalues, J. Amer. Math. Soc. 2(1989), 187~196.

[5] Li, W.-C. W., Character sums and abelian Ramanujan graphs, J. Number Theory, 41(1992), 199~217.

[6] ——, Number Theory with Applications, World Scientific, 1996.

[7] Lubotzky, A. , Phillips, R. , Sarnak, P. , Ramanujan graphs, Combinatorica 8(1988), 261～277.

[8] Margulis, G. , Explicit construction of concentrators, Problems of Information Transmission 9(1975), 325～332.

[9] ——, Explicit group theoretic constructions of combinatorial schemes and their application to the design of expanders and concentrators, J. Prob. of Info. Trans. (1988), 39～46.

[10] Mckay, B. D. , The expected eigenvalue distribution of a large regular graph, Linear Alg. and Its Appli. 40(1981), 203～216.

[11] Morgenstern, M. , Existence and explicit constructions of $q+1$ regular Ramanujan graphs for evary prime power q, J. Comb. Theory, (B)62(1994), 44～62.

[12] Nilli, A. , On the second eigenvalue of a graph, Disc. Math. 91(1991), 207～210.

[13] Pizer, A. , Ramanujan graphs and Hecke operators, Amer. Math. Soc. Bull. 23(1990), 127～137.

[14] Serre, J. -P. , Trees, Springer-Verlag, 1980.

[15] Lubotzky, A. , Discrete Groups, Expanding Graphs and Invariant Measures, Progress in Math. 125, Birkhausen, 1994.

关于图的 zeta 函数.

[1] Bass, H. , The Ihara-Selberg zeta function of a tree lattice, Internat. J. Math. 3(1992), 717～797.

[2] Foata, D. , Zeilberger, D. , A combinatorial proof of Bass's evaluations of the Ihara-Selberg zeta function for graphs, Trans. Amer. Math. Soc. 351(1999), 2257～2274.

[3] Hashimoto, Zeta functions of finite graphs and representations of p-adic groups, Adv. Stud. Pure Math. 15(1989), 211～280, Academic Press.

[4] ——, Artin-type L-functions and the density theorem for prime cycles on finite graphs, Internat. J. Math. 3(1992), 809～826.

[5] Ihara, Y., On discrete subgroups of the 2×2 projective linear group over p-adic fields, J. Math. Soc. Japan 18(1966), 219～235.

[6] Stark, H. M., Terras, A. A., Zeta function of finite graphs and coverings, Adv. in Math. 121(1996), 124～165.

[7] Sunada, T., L-functions in geometry and some applications, Lecture Notes in Math. 1201, pp. 266～284, Springer-Verlag, 1986.

[8] ——, Fundamental groups and Laplacians, Lecture Notes in Math. 1339, pp. 248～277, Springer-Verlag, 1988.

[9] Venkov, A. B., Nikitin, A. M., The Selberg trace formula, Ramanujan graphs and some problems of mathematical physics, Petersburg Math. J. 5(1994), 419～484.

结　语

1. 本书概括地介绍了代数数论的起源、发展和现状. 由于篇幅和作者的学识所限，不可能叙述得非常全面，甚至对代数数论的一些重要理论发展和应用都没有深入地触及. 比如说，在理论方面，我们没有介绍算术几何中的 Arakelov 理论（用现代微分几何研究整体域上的代数簇），也没有介绍代数数论和代数 K-理论的联系. 这些研究与书中介绍的内容具有同等的重要性. 在应用方面，书中主要介绍了代数数论在通信领域方面的应用，没有涉及在其他领域的渗透. 比如说，我们没有谈及 J. Silverman 等人用局部域和类域论工具研究动力系统的轨道特性，也没有涉及 Witten 用 Adele 工具研究弦理论和物理学家们用 p-adic 分析研究量子场论所作的尝试.

2. 韦依（A. Weil）晚年致力于数论史的研究，对于数学史研究的目的和服务对象发表过许多看法. 他认为数学史不仅是把各种数学事实按照时间、国家、学科和作者加以考究和分类. 他同意莱布尼兹（Leibniz）的观点："历史不仅要弄清楚每个人的功绩，而且要以生动的例子揭示出发现的艺术和方法."数学史的服务对象可以是受过教育的外行（educated layman），也可以是为历史学家们. 他认为数学史最重要的功能是为年青人展示历史上第一流数学家们如何做数学的"生动例子". 数学思想是数学史的主要研究对象. 每个年青人不仅向同时代的数学家，也应向历史上

的杰出数学家了解和学习他们从事数学研究的战术和战略. 所谓战术即指“学者们每天所使用的工具”，而“数学战略是长期目标，是识别主要问题、寻求突破口、攻克难点、预见未来目标和研究计划的艺术. 这需要对广泛的领域和思想的长期演化有深刻的理解.” 韦依认为这是研究数学史和学习数学发展历史的最主要目的.

韦依举例说，17 世纪的英国历史学家热衷于介绍牛顿是造币厂的厂主，或是某位贵妇人的叔叔，觉得这些比牛顿做为数学家更有兴趣. 公元 7 世纪印度人 Bhaskara 给出某些具体的 Pell 方程整数解，17 世纪费马等人又加以研究. 有些人甚至考证这些工作起源于希腊人(包括阿基米德). 但是没有人真正说清 Bhaskara 和 17 世纪数学家关于 Pell 方程工作的联系. 数学家的生平、个性、和生活环境也是数学史的研究内容，因为这对数学家的学术活动是有影响的. 韦依本人的数论史著作中就有许多这方面的例子.

3. 代数数论的历史对于探讨数学发展的动力，数学的功能，数学教育和年青数学人材的培养等许多方面都有很好的启发性和借鉴性. 数论起源于生产和生活实践，由古希腊人把它提升到理性的高度. 在研究不定方程整数解的过程中，有理数域被扩大为二次域、分圆域以至任意代数数域，有理整数被推广为代数整数，整数环上的惟一因子分解特性被推广成任意代数整数环的素理想惟一因子分解特性，并建立了理想类群和类数概念，初等数论扩展成内容丰富的“高级算术”——经典代数数论. 19 世纪中期之后，黎曼 zeta 函数到代数数域的推广把解析方法与代数方法结合起来，赋值论和局部化方法又把数论与几何结合起来. 20 世纪中期之后，现代数论与纯粹数学的诸多领域融合交织在一起，这种多个领域研究思想和方法的融合和渗透产生出一系列重大的理论研究成果. 费马猜想于 1994 年被证明是这些成果的光辉代表，这项成果吸取了近年来数论和相关领域许多重要的思想和方法，是体现现代数论和算术代数几何进步的结晶. 费马猜想被证明并没

有结束数论的发展. 1994 年之后数论的重大发展有：

(1) L. Merel 于 1996 年证明了下面的重要猜想 (Inv. Math. 124 (1996), 437～449)：对每个正整数 d，均存在只依赖于 d 的正整数 B_d，使得每个 d 次代数数域 K 和 K 上每条椭圆曲线 E，E 的 K-有理点群 $E(K)$ 的 torsion 部分 $E(K)_t$ 都是阶数 $\leqslant B_d$ 的有限阿贝尔群. 在此之前，B. Mazur (1977 年) 和 S. Kamienny 分别对有理数域和二次域证明了这个猜想，并且决定出群 $E(K)_t$ 的所有可能的结构.

(2) A. Wiles 对于半稳定的椭圆曲线证明了 STW 猜想. 并由此推出费马猜想的正确性. 1999 年，C. Breuil，B. Conrad，F. Diamond 和 R. Taylor 宣布完全证明了 SWT 猜想. 见 H. Darmon, A proof of the full Shimura-Taniyama-Weil conjecture is announced, Notes AMS46 (1999), 1397～1401.

(3) 1998 年，Michael Harris 和 Richard Taylor 证明了局部数域情形下 GL_n 的 (局部) 朗兰兹猜想，不久 Guy Henniart 又给了较简单的证明. 另一方面，对于函数域的情形，Laurent Lafforgue 把 Drinfeld 关于 $n=2$ 情形的获费尔兹奖工作成功地加以推广，对于函数域和任何 $n\geqslant 1$，证明了关于 GL_n 的 (整体) 朗兰兹猜想. 这无疑又是代数数论发展的一个里程碑. 注意 R. Taylor 即是 A. Wiles 证明费马猜想的合作者. 见 J. Rogawski, The Nonabelian Reciprocity Law for Local Fields, Notes AMS 47 (2000) 35～41.

数论在理论研究中取得辉煌的成就，与此同时，它在实际领域得到一系列重大的应用. 深刻的理论成果在实际领域中的这些应用不仅令人惊叹，而且常常被人们认为是“不可预见”的. 计算数论队伍 (由数论学家、通信专家和计算机专家相结合) 在世界各地相继出现和壮大，不仅为数论研究增添了活力，而且促进技术水平的根本性的提高. 在数学发展史中，我们看到非欧几何对狭义相对论所起的作用，看到力学和电磁学对微积分和数理方程的产生和发展所起的作用，微分几何与拓朴学对理论物理所起

的作用. 在 21 世纪，数论和代数几何对于信息领域将起到类似的互动作用.

我们生活在数学蓬勃发展、数学向其他领域深层次地扩展，数学本身不断融合的时代，尽快把现代数学教给年轻人，努力培养具有数学思想和能力的新一代（这些人或者从事数学研究，或者在其他领域工作），是数学教育的重要任务. 营造讨论与合作的气氛和活跃的学术交流环境，会培养出高水平的一代数学人材.

[General Information]
书名=代数数论简史
作者=冯克勤著
页数=276
SS号=11134138
DX号=
出版日期=2002年07月第1版
出版社=湖南教育出版社